In the hunt for supernovae, this amateur is a star p. 20

Expansion microscopy reveals secrets of transcription pp. 26 & 92

Silent flagellins win NOSTER & *Science* Microbiome Prize p. 37


# Science

$15
7 JULY 2023
science.org

AAAS

## TIGHTENING BOUNDS

Molecular ion spectroscopy constrains electron's dipole moment pp. 28 & 46

EDITORIAL

# Museum ethics—more than lip service

In the past few years, museums around the world have started to grapple with questions about the origins and ethics of their collections. This includes the acquisition and maintenance of natural history specimens. As museums examine their missions and processes, it seemed like a good time to talk to Sean Decatur, the new president of the American Museum of Natural History (AMNH) in New York City. He spoke with me (see the full conversation here) about the museum's research and how collaborations between museums and partner nations should aspire to assemble collections that responsibly disseminate information about human cultures, the natural world, and the Universe.

Last December, Decatur transitioned from university administration (as president of Kenyon College in Ohio for 9 years) to museum leadership. He makes a strong case that natural history collections are critical for science. The AMNH's roughly 34 million specimens and objects, especially the biological collection, "tells the story of life on the planet," he said. "You can really follow changes that have happened in species that might have been collected from the same geography over the course of more than a century now." But Decatur is sensitive to how museums must responsibly build and keep such massive collections. He acknowledged that standards have changed over time "not only in terms of agreements about ownership of items being collected, but collaborations that involve the engagement of students and faculty from the partner nations and institutions where we're collecting." He boils it down to creating "a true two-way collaboration with students moving in both directions along the development of a project."

Determining the proper home for objects from Indigenous groups in the United States is governed by the Native American Graves Protection and Repatriation Act of 1990. I asked Decatur how the museum views compliance with these regulations when questions are raised about whether objects or specimens should be in New York. For items that belong to North American Indigenous groups, he emphasized that a clear process exists for repatriation and that the museum has resources to work with Indigenous groups who claim ownership of objects that are in the museum's collection. But he also wants to make sure that commitments are "more than lip service" by ensuring that the museum returns items that are not now, and were not in the past, collected under terms that do not meet today's ethical standards. Moreover, Decatur is focused on building fulsome partnerships devoted to healing and moving forward from the past. "There's more to returning items as a repatriation process than just putting them in the mail," he said.

> **"There's more to returning items as a repatriation process than just putting them in the mail."**
> **Sean Decatur**, American Museum of Natural History

Earlier this year in May, the AMNH opened the Richard Gilder Center for Science, Education, and Innovation, which includes classrooms and programs aimed at helping students—elementary grades through high school—understand that science is a process and not just a collection of facts. Decatur believes strongly in the ability of museum collections and exhibitions to provide a tangible connection to textbooks. This relationship "can give confidence in what science has to tell us about the world," he said. "You can't really develop that confidence unless there's an understanding of what the process is like. The museum offers a great opportunity for that, partially because there are objects that tend to capture excitement and wonder." He admitted, "I still walk through the insectarium and look at the range of beetles that are in the collection and it just inspires a sense of wonder about the natural world."

Museums must take on the important work of self-auditing and evaluating their conceptualization of what is collected and how, the stories their exhibits tell, the perspectives and processes involved, and the larger research and educational roles that they serve. Sean Decatur knows that these overhauls start with decision-makers like him.

**–H. Holden Thorp**

**H. Holden Thorp**
Editor-in-Chief, *Science* journals.
hthorp@aaas.org;
@hholdenthorp

PHOTO: CAMERON DAVIDSON

10.1126/science.adj4713

# NEWS

## IN BRIEF

*Edited by* **Jeffrey Brainard**

New AI techniques could help better predict cyclones such as 2018's Typhoon Yutu, which intensified rapidly before hitting Tinian and Saipan.

METEOROLOGY

# Artificial intelligence forecasts the weather in a flash

A model using artificial intelligence (AI) forecasts global weather as accurately, and more than 10,000 times faster, than the best system in use, researchers report. The conventional tool, run by the European Centre for Medium-Range Weather Forecasts, is computationally intensive, requiring hours of supercomputer calculations to produce a 10-day forecast. The new AI-based model—named Pangu-Weather and developed by Huawei, the Chinese tech giant—improves on previous AI-powered models by simulating weather at different altitudes and forecasting tropical cyclones, with results reliable out to 10 days, researchers say. The research team trained the model on 39 years of historical weather data; the system is untested yet using real-time observational data. Another AI-based weather model, GraphCast, described by Google DeepMind in a December 2022 preprint, also outperformed the European system.

## CRISPR relatives found all over

GENETICS | The revolutionary invention of the gene editor called CRISPR converted an immune mechanism used by single-celled organisms into one of the mostly widely used tools in biology. Now, two groups have discovered that many multicelled organisms, including fungi, plants, and animals, carry genes for a DNA-cleaving enzyme similar to one that prokaryotes use in CRISPR—a discovery that may lead to powerful improvements in gene editing. In a bioRxiv preprint and in *Nature,* the research teams, both of which are affiliated with the Massachusetts Institute of Technology, also reported that, as with CRISPR, they could program RNA guides coupled to these enzymes to shuttle them to specific locations in human cells to precisely edit genes. The researchers speculate that eukaryotic cells may have gained the newly identified editing genes from transposable elements—so-called jumping genes—they received from bacteria. Their function in eukaryotes is unknown.

## Nobelist claims antimale bias

COMMUNITY | Organizers of a conference that brings together Nobel laureates and early-career researchers are investigating whether one of the prize winners violated the meeting's code of conduct last week when he disparaged its focus on scientists' diversity during a session about structural biology. At the Lindau Nobel Laureate Meeting, Kurt Wüthrich, who was awarded a Nobel Prize in Chemistry in 2002, also referred to "discrimination against" men

researchers. That comment drew pointed criticism from an early-career researcher in the audience: "There might be individual discrimination towards men, but it is nothing compared to the systematic and structural discrimination that women have to face, especially in the STEM fields," she said, eliciting applause from some in the audience. (She did not identify herself.) The code of conduct bars comments that "reinforce social structures of domination related to … gender." To date, 60 women have been awarded Nobel Prizes, compared with 892 men. Just five of 39 invited speakers at this year's Lindau meeting were women.

## Ancient DNA expert finds new job

**WORKPLACE** | Alan Cooper, a noted specialist in ancient DNA dismissed by the University of Adelaide in 2019 for "serious misconduct," has been appointed a professor at Australia's Charles Sturt University (CSU). Cooper, who once ran Adelaide's Australian Centre for Ancient DNA, faced allegations of bullying members of his institute, particularly students. Cooper will be affiliated with CSU's Gulbali Institute in Albury-Wodonga. In a statement, CSU confirmed Cooper's employment, first reported by *Nature*, but declined to give details, citing privacy obligations. The statement said all academic staff "are subject to probation and supervisory requirements" and that the university "is firmly committed to ensuring a positive workplace culture." Cooper did not respond to a request for comment before *Science* went to press.

## Green hydrogen method advances

**MATERIALS SCIENCE** | U.S. and European researchers last week reported making big improvements to a green way for efficiently making hydrogen, a carbon-free fuel. The team varied the recipe for solar cells that can power catalysts to split water molecules into oxygen and hydrogen. Researchers previously made the solar cells from materials including so-called III-V semiconductors, which are highly efficient. But these devices corrode quickly in water, rendering them impractical. Now, the team has replaced the III-Vs with devices made from layers of silicon and perovskite, a cheaper material that blends organic and inorganic compounds, and protected both with a corrosion-resistant coating. The devices can convert 20% of solar energy into hydrogen gas and survive for more than 100 hours of operation, beating the III-Vs on both measures, the researchers reported in *Nature Communications*. But they need even longer lifetimes to become practical.

An Awá woman holds hunting bows and arrows in Brazil's Caru Indigenous Territory in 2017.

PHOTO: SCOTT WALLACE/GETTY IMAGES

**ANTHROPOLOGY**

## Worldwide survey challenges idea of 'Man the Hunter'

Women have hunted in nearly 80% of forager societies surveyed in a first-of-its-kind global review that contradicts the once-prevailing view that men exclusively hunt while women gather. In the 20th century, influential anthropologists championed the notion that hunting and meat consumption drove humanity's key evolutionary changes, including big brains, with men relying on their strength to hunt while women stayed near camp, gathering plants and tending children. But the survey data refute this. A research team examined ethnographic reports of 63 forager groups spanning the late 1800s to 2010s and found that women hunted in 50 of them. The study, published last week in *PLOS ONE*, found both men and women took a variety of prey, although they sometimes practiced different hunting methods.

## Do averages hamper innovation?

**PEER REVIEW** | Funding agencies might finance a larger number of risky research ideas if they ditched the common practice of averaging the quality scores provided by reviewers of grant applications, a study says. It found that biomedical scientists participating in a mock review were willing to recommend funding proposals with slightly lower average scores if they drew high scores from some reviewers. "It appears that our experts valued healthy disagreement" among reviewers over a proposal's value, the research team reports in a National Bureau of Economic Research working paper published in June. They asked 563 scientists funded by the U.S. National Institutes of Health, which typically uses averaged scores to rank proposals, to make funding decisions based on the scores, title, and abstract of grant proposals. Although the participants chose some studies with high average scores, they also liked lower-scored proposals with high variation in individual reviewer scores—an approach that changed the ranking of 58% of proposals compared to a ranking system based only on average scores.

## Giant rockslides cut Himalayas

**GEOLOGY** | Nearly 1000 years ago, a giant rockslide in central Nepal collapsed the summit of a mountain likely taller than 8000 meters—a phenomenon that may have occurred on other peaks across the Himalayas, taming the world's highest mountain range, researchers report. The peak probably lost several hundred meters of elevation and 23 cubic kilometers of rock—a volume equivalent to that of 22,000 Empire State Buildings—making it the Himalayas's largest known rockslide, according to the study this week in *Nature*. High-altitude permafrost may have helped the range's tall peaks resist erosion by weather until the underlying rock gave way in a massive failure, perhaps triggered by an earthquake or local warming. The collapses, the authors conclude, are the ultimate control on the Himalayas's towering heights.

The Supreme Court ruling brought out critics like two Harvard students (right) as well as supporters.

EDUCATION

# After race ban, other routes to diversity beckon

## Equity advocates see paths to increasing minority representation—but say it won't be easy

*By* **Jeffrey Mervis**

Last week's decision by the U.S. Supreme Court to ban the use of race in undergraduate admissions was a serious setback, say advocates of increasing minority participation in science and engineering. But they are hoping that one paragraph in the court's ruling will accelerate adoption of alternative practices meant to accomplish the same goal: an education system that better meets the needs of the country's diverse population.

"I'm not sure it's 100% bad," says mathematician Richard Tapia, a winner of the National Medal of Science for both his research and his work mentoring students from groups underrepresented in science. "It'll force us to change our approach to admissions. And I feel pretty confident that we can meet the challenge."

Jamie Lewis Keith, an attorney with consulting firm EducationCounsel, is telling her university clients not to panic. "Do not retreat from your equity mission," says Keith, who also works with SEA Change, an AAAS initiative to improve campus diversity (AAAS publishes *Science*). "Reach out to students of color—reassure them that they are welcome, belong, and will be supported at your institution."

In overlapping 6-2 and 6-3 rulings on 29 June, the high court said the race-conscious undergraduate admissions practices of Harvard University and the University of North Carolina, respectively, violated the U.S. Constitution, specifically the equal protection clause of the 14th Amendment. "Universities wishing to discriminate based on race must justify a compelling state interest," wrote Justice Clarence Thomas in a concurring opinion on the case, brought by the advocacy group Students for Fair Admissions. "I highly doubt any will be able to do so."

In a sharp dissent, Justice Sonia Sotomayor said the decision was a step in the wrong direction. "It rolls back decades of precedent and momentous progress," she wrote. "Ignoring race will not equalize a society that is racially unequal," she added.

The court's decision effectively ends 6 decades of affirmative action in admissions by colleges and universities. Modified over time by numerous legal challenges, the use of affirmative action in its current form draws on extensive research showing that increasing campus diversity improves the quality of education that students receive. But its status was always precarious, and the court's conservative majority had repeatedly signaled its intentions to eliminate race as a criterion in admissions. "We always knew this day would come," Tapia says.

In practice, the ruling only applies to those 200 or so most selective U.S. universities that receive applications from many more students than they can admit. But those are the same institutions that train most academic researchers, whose role includes teaching the next generation of those working in science, technology, engineering, and math (STEM) fields.

"Graduate STEM programs recruit from the most selective universities," says Stella Flores, who studies education leadership and public policy at the University of Texas at Austin. "So they will take the biggest hit if those universities back away" from a commitment to diversity in undergraduate admissions.

Advocates for broadening participation in STEM say they are encouraged by one brief passage in the majority opinion written by Chief Justice John Roberts. It says a student's race may be relevant in the admissions process if an applicant discusses how it shaped them as an individual. Roberts offers the example of a student essay that discusses "overcoming racial discrimination" to demonstrate "courage and determination" or how an applicant's "heritage or culture" prompted

them "to assume a leadership role."

Keith says those experiences, showing resilience or what has been called grit, can be taken into account so long as the demonstrated character qualities contribute to the educational mission of the institution. That's especially important in STEM fields, she says, where many minority students haven't had the same opportunities as their white peers to take the needed prerequisite courses and begin to self-identify as a potential scientist.

Tapia, who describes his career path as going "from the barrios of Los Angeles to the White House," sees the ruling as an invitation for research universities to go beyond the traditional reliance on grades and standardized test scores in assessing a student's potential. "We need to make sure that the person that we accept has a high probability of succeeding," he says. The need to assess students' experiences as well as their grades and scores has implications for the admissions committees themselves, he says. "To do that, you need to have the appropriate background. I understand what that takes because I've been there."

Ruth Simmons, a former president of Brown University who just retired from the presidency of Prairie View A&M University, a historically Black institution, agrees. "We use proxies, like test scores or where a person was educated, instead of looking at the whole person," she says. "And this decision, at its best, may get us out of that lazy way of assessing people. So I'm excited about the possibilities."

Tapia emphasizes that admissions aren't the only factor in diversity. Retaining those students once they are admitted is equally important.

"You don't wait for a minority student to come and ask you for help. That's too late," he says. "They don't want to admit that they're not doing so well. So I reach out to those who I find out are struggling."

University administrators say they are still trying to digest the ruling and looking for guidance. They may get some soon from the White House, which announced it would convene a national summit on "expanding access to educational opportunity." The gathering will be followed in September by a report from the Department of Education on how to use "measures of adversity … to build inclusive, diverse student bodies."

Applying those measures will require what Flores calls authentic recruiting, in which faculty and senior administrators are trained to recognize the barriers that students have faced in seeking a quality education. "It's going to take real courage and conviction," she adds. "And a lot more money." ■

SCIENTIFIC COMMUNITY

# Winning a Nobel Prize may be bad for your productivity

## Fame and time commitments that come with the honor distract from continued research

*By* **David Shultz**

For many scientists, there's no greater achievement than winning a Nobel Prize. The award recognizes breakthroughs that have broadened our understanding of reality and changed the world. It also might be a productivity killer. According to a working paper recently published by the National Bureau of Economic Research, scientists' output tends to fall off a cliff after they've won the prize.

Researchers behind the study analyzed data on Nobel Prize winners in physiology and medicine between 1950 and 2010, charting how three factors changed after their awards: the number of papers they published; the impact of those papers, based on how often they were cited; and how novel their ideas were. The authors quantified novelty by using a computer program to identify distinct scientific ideas in the papers, then assigning each idea an age based on when it first appeared in the U.S. National Library of Medicine's Unified Medical Language System. Papers with "younger" ideas relative to their publication dates were given higher novelty scores.

> **"They're not just scientists anymore, but public intellectuals."**
> **Jayanta Bhattacharya**, Stanford University

Because the prize often goes to scientists late in their careers, when output typically declines, the study compared Nobelists with age-matched winners of another prestigious medical prize, the Lasker Award. "The matching strategy is sound," says Kirk Doran, a social scientist at the University of Notre Dame. "Overall, the results are quite convincing."

Before winning the prize, future Nobel laureates published more frequently than their colleagues who eventually won the Lasker Award, and they also published more-novel papers and garnered more citations. After winning a Nobel, however, the trend flipped: Nobel winners, on average, saw declines in productivity, novelty, and citations, falling even with the Lasker winners or below them. Future Nobel winners published about one study more per year than future Lasker winners in the 10 years leading up to the prize. In the 10 years after winning, however, the Lasker group published about one more study annually than did their Nobel laureate peers. The switch was driven almost entirely by declines in the Nobel group.

"I was actually surprised by the shift in novelty," says the working paper's first author, Jayanta Bhattacharya, a health economist and epidemiologist at Stanford University. "I expected the shifts in productivity because of the [new] demands on their time, but the shift towards less novel work is a little bit of a surprise to me."

Doran, who has studied the career effects of winning a Fields Medal—the Nobel's equivalent in math—says the results gel with his own findings, which show that mathematicians also appear to suffer decreases in productivity following a win.

Although Bhattacharya and colleagues stress that their analysis does not show a causal link between winning the Nobel and a drop in productivity, Doran notes that winning such a prestigious award is life-changing. Laureates may be overwhelmed with speaking engagements, media interviews, or book deals, all of which eat into the time and energy available to do original science. The Lasker, although highly prestigious, does not carry the same household recognition.

"There's just a big change in the patterns that you see after the Nobel Prize, that indicates a pretty fundamental shift in their lives," Bhattacharya says. "They're not just scientists anymore, but public intellectuals."

Given that, do the prize's benefits outweigh the potential for lost research? Bhattacharya doesn't want to scrap the Nobel Prize, but suggests saving the recognition for scientists later in their career, to prevent interrupting our most brilliant minds in their prime. "Awarding Nobels as career achievement awards would still achieve a lot of the benefits of the Nobel while eliminating some of the harms," he says. ■

David Shultz is a journalist in Santa Barbara, California.

them "to assume a leadership role."

Keith says those experiences, showing resilience or what has been called grit, can be taken into account so long as the demonstrated character qualities contribute to the educational mission of the institution. That's especially important in STEM fields, she says, where many minority students haven't had the same opportunities as their white peers to take the needed prerequisite courses and begin to self-identify as a potential scientist.

Tapia, who describes his career path as going "from the barrios of Los Angeles to the White House," sees the ruling as an invitation for research universities to go beyond the traditional reliance on grades and standardized test scores in assessing a student's potential. "We need to make sure that the person that we accept has a high probability of succeeding," he says. The need to assess students' experiences as well as their grades and scores has implications for the admissions committees themselves, he says. "To do that, you need to have the appropriate background. I understand what that takes because I've been there."

Ruth Simmons, a former president of Brown University who just retired from the presidency of Prairie View A&M University, a historically Black institution, agrees. "We use proxies, like test scores or where a person was educated, instead of looking at the whole person," she says. "And this decision, at its best, may get us out of that lazy way of assessing people. So I'm excited about the possibilities."

Tapia emphasizes that admissions aren't the only factor in diversity. Retaining those students once they are admitted is equally important.

"You don't wait for a minority student to come and ask you for help. That's too late," he says. "They don't want to admit that they're not doing so well. So I reach out to those who I find out are struggling."

University administrators say they are still trying to digest the ruling and looking for guidance. They may get some soon from the White House, which announced it would convene a national summit on "expanding access to educational opportunity." The gathering will be followed in September by a report from the Department of Education on how to use "measures of adversity … to build inclusive, diverse student bodies."

Applying those measures will require what Flores calls authentic recruiting, in which faculty and senior administrators are trained to recognize the barriers that students have faced in seeking a quality education. "It's going to take real courage and conviction," she adds. "And a lot more money." ■

SCIENTIFIC COMMUNITY

# Winning a Nobel Prize may be bad for your productivity

## Fame and time commitments that come with the honor distract from continued research

*By* **David Shultz**

For many scientists, there's no greater achievement than winning a Nobel Prize. The award recognizes breakthroughs that have broadened our understanding of reality and changed the world. It also might be a productivity killer. According to a working paper recently published by the National Bureau of Economic Research, scientists' output tends to fall off a cliff after they've won the prize.

Researchers behind the study analyzed data on Nobel Prize winners in physiology and medicine between 1950 and 2010, charting how three factors changed after their awards: the number of papers they published; the impact of those papers, based on how often they were cited; and how novel their ideas were. The authors quantified novelty by using a computer program to identify distinct scientific ideas in the papers, then assigning each idea an age based on when it first appeared in the U.S. National Library of Medicine's Unified Medical Language System. Papers with "younger" ideas relative to their publication dates were given higher novelty scores.

> **"They're not just scientists anymore, but public intellectuals."**
> **Jayanta Bhattacharya**, Stanford University

Because the prize often goes to scientists late in their careers, when output typically declines, the study compared Nobelists with age-matched winners of another prestigious medical prize, the Lasker Award. "The matching strategy is sound," says Kirk Doran, a social scientist at the University of Notre Dame. "Overall, the results are quite convincing."

Before winning the prize, future Nobel laureates published more frequently than their colleagues who eventually won the Lasker Award, and they also published more-novel papers and garnered more citations. After winning a Nobel, however, the trend flipped: Nobel winners, on average, saw declines in productivity, novelty, and citations, falling even with the Lasker winners or below them. Future Nobel winners published about one study more per year than future Lasker winners in the 10 years leading up to the prize. In the 10 years after winning, however, the Lasker group published about one more study annually than did their Nobel laureate peers. The switch was driven almost entirely by declines in the Nobel group.

"I was actually surprised by the shift in novelty," says the working paper's first author, Jayanta Bhattacharya, a health economist and epidemiologist at Stanford University. "I expected the shifts in productivity because of the [new] demands on their time, but the shift towards less novel work is a little bit of a surprise to me."

Doran, who has studied the career effects of winning a Fields Medal—the Nobel's equivalent in math—says the results gel with his own findings, which show that mathematicians also appear to suffer decreases in productivity following a win.

Although Bhattacharya and colleagues stress that their analysis does not show a causal link between winning the Nobel and a drop in productivity, Doran notes that winning such a prestigious award is life-changing. Laureates may be overwhelmed with speaking engagements, media interviews, or book deals, all of which eat into the time and energy available to do original science. The Lasker, although highly prestigious, does not carry the same household recognition.

"There's just a big change in the patterns that you see after the Nobel Prize, that indicates a pretty fundamental shift in their lives," Bhattacharya says. "They're not just scientists anymore, but public intellectuals."

Given that, do the prize's benefits outweigh the potential for lost research? Bhattacharya doesn't want to scrap the Nobel Prize, but suggests saving the recognition for scientists later in their career, to prevent interrupting our most brilliant minds in their prime. "Awarding Nobels as career achievement awards would still achieve a lot of the benefits of the Nobel while eliminating some of the harms," he says. ■

David Shultz is a journalist in Santa Barbara, California.

PALEONTOLOGY

# Rare fossil implies deep roots for vertebrates

## Half-billion-year-old tunicate from western Utah "looks like it died yesterday"

*By* **Amanda Heidt**

In 2019, a finger-size fossil landed on the desk of Karma Nanglu, a paleontologist at Harvard University who specializes in the Cambrian and Ordovician periods, when many of today's animal forms made their entrance. The specimen had sat for years in the drawer of a Salt Lake City museum; its finders, who had pulled it from a fossil-rich layer of Cambrian limestone in western Utah, thought it might be a sea squirt or tunicate—a marine invertebrate that shares a distant ancestor with all vertebrates. Nanglu was excited but cautious at the prospect of a very ancient tunicate: "That's a group for which there is essentially no fossil record for the entire 500 million years of recorded history."

An ancient tunicate (fossil, left; reconstruction, right) had a saclike body and twin siphons, much like tunicates today.

Now, in a paper published in *Nature Communications* this week, Nanglu and his coauthors report that the exquisitely-preserved 500-million-year-old fossil is a dead ringer for some tunicates today, with two siphons to filter organic particles from the water and complex musculature controlling the siphons. "It looks like a tunicate that died yesterday and just happened to fall down on some rock," says Nicholas Treen, a developmental biologist at Princeton University who wasn't involved in the work. The discovery offers clues to the timing and development of early tunicates and could even push back the date for the origin of tunicates' sister group, the vertebrates, including humans.

Today, some 3000 species of tunicate live in almost every habitat of the oceans. Most have a two-part life cycle, including a free-swimming, tadpolelike larva that settles and metamorphoses into a stationary adult. Tunicate larvae have a notochord, the precursor to a spinal column—a defining trait of the group called chordates, which includes all vertebrates. But only a handful of tunicate fossils exist, for reasons paleontologists can't fully explain. The best-known, 520-million-year-old *Shankouclava anningense*, has proven controversial: It's tunicate-ish in that it has a siphon and seemed to filter feed, but it doesn't resemble anything alive today.

The new specimen, however, with details of its soft body preserved, is unambiguously a tunicate. "It's really kind of glorious, the picture of this fossil," Treen says. Nanglu and colleagues named it *Megasiphon thylakos* for its large siphons and saclike body (*thylakos* is from the Greek word for pouch). Modern tunicates use these siphons to help filter-feed and have bands of longitudinal muscles running down their bodies as well as circular muscles around the siphons to control them. The details of *Megasiphon*'s musculature look almost identical to those of *Ciona intestinalis*, a well-studied modern tunicate, Nanglu says.

As a tunicate matures, its siphon muscles develop from the same cells that form cardiac tissue in modern vertebrates. The muscles' presence in *Megasiphon* means it may have already had something like a heart, even though its internal structures aren't preserved. "Since you can see these atrial siphon muscles, you can almost take it for granted that there is a beating, vertebrate-like heart inside this organism," Treen says.

The fossil addresses a long-unresolved question in early chordate evolution: whether the common ancestor to all tunicates was a free-swimming organism or rooted to the bottom. *Megasiphon*, with its resemblance to living, sessile tunicates, strongly supports the latter hypothesis. The find suggests that tunicates' two-part life history and ability to metamorphose is an ancestral characteristic of the group.

The existence of such a complete and recognizable tunicate at this time could also push back the origin of vertebrates, now thought to be at about 450 million years ago. "If a tunicate body plan was already established [at 500 million years ago], and the tunicates are actually the sister group of vertebrates, then vertebrates may be older as well," says William Jeffery, a developmental biologist at the University of Maryland, College Park, who was not involved in the work. Other fossils offer some support to this idea: *Megasiphon* likely lived alongside a swimming, fishlike animal named *Metaspriggina*, dated to as old as 505 million years ago. *Metaspriggina* had eyes and blocks of muscle, and probably a notochord. So, by about 500 million years ago, both it and *Megasiphon* already have many features considered as vertebrate hallmarks, suggesting an early origin for the group.

Still, Nanglu says existing fossils aren't enough to change the story of vertebrate evolution. "Humans are naturally pulled to this question of origins, and this helps develop that story, but we need more fossil evidence," he says. "But … answering these longstanding questions, that's totally why we're in it to begin with." ■

Amanda Heidt is a science journalist in Moab, Utah.

Religious texts are a focus of instruction in Israel schools for the ultra-Orthodox, especially for boys.

PHOTO: RONEN ZVULUN/REUTERS

EDUCATION

# Israel's science-free schools for ultra-Orthodox Jews draw fire

## Researchers criticize approval of new funding for religious schools without requirements for broader curriculum

*By* **Michele Chabin**, *in Jerusalem*

The Israeli parliament's recent decision to give a hefty budget increase to religious schools that often do not teach science and math is drawing criticism from the nation's research community. The move threatens to leave a growing share of Israel's young people without the knowledge and skills needed to thrive in an increasingly competitive global economy, they warn.

"Hundreds of thousands of kids are getting an education for life in the Middle Ages," says biochemist and Nobel laureate Aaron Ciechanover of the Technion Israel Institute of Technology. "I won't say that what they're learning is useless, but it won't make them part of the Israeli economy or prepare them to study at university."

At issue are hundreds of elementary and secondary schools that serve Israel's ultra-Orthodox, or Haredi, Jewish community. Haredim comprise roughly 13% of Israel's population, but Haredi schools enroll nearly 26% of Jewish students and 20% of all Israeli students. The schools, which receive public funding but often are not subject to government regulation, enrolled about 373,000 students during the 2021–22 school year.

Haredi schools typically prepare boys for lifelong study of Jewish religious texts and law, rather than employment. By seventh grade, most of the curriculum focuses on religious content. In 2020, for example, 84% of boys in Haredi high schools studied no secular subjects, according to political scientist Gilad Malach of the Israel Democracy Institute. Haredi girls, who are expected to work and support their families once they marry, are taught a state-mandated core curriculum that includes secular subjects, but the lessons are often less advanced than those given to non-Haredi students.

The quality of Haredi education has long been a sensitive issue in Israel. Supporters of Haredi schools say they are intellectually rigorous and help meet a sacred obligation to strengthen the Jewish character of the world's only Jewish-majority country. But critics argue they contribute to high poverty and unemployment rates in the Haredi community, which demographers forecast could constitute 25% of Israel's population by 2050. Roughly half of Haredi men are jobless.

Earlier this year, government funding for Haredi schools became a political flash point after Prime Minister Benjamin Netanyahu's governing coalition, which draws critical support from conservative religious and nationalist parties, proposed a $333 million spending boost for the schools in its 2023–24 budget. In the past, Netanyahu supported linking public funding for Haredi schools to their willingness to teach secular subjects, including science, math, and English. But the budget proposal included no such conditions because of opposition from key members of Netanyahu's fragile coalition.

As parliament prepared to vote on the budget in May, some academic researchers urged it to reject the no-strings-attached funding. "Research clearly shows that children who do not study a full core curriculum have considerable difficulty in overcoming this disadvantage as adults," stated one open letter signed by nearly 300 economists. "The average proficiency of Israeli students in math, science and reading already places them below the average score in every other developed country," they noted, arguing that Netanyahu's government was pursuing policies "on the basis of short-term political considerations" that "are putting Israel on the path to becoming a third world economy." Such warnings, however, failed to prevent lawmakers from approving the budget on 24 May, with Netanyahu winning support from all 64 members of his coalition.

If Netanyahu's uneasy governing alliance ultimately fractures, a new government could revisit the spending plan. In the meantime, some science and technology advocates are working to strengthen programs that provide technical training to graduates of Haredi schools. The Israel Democracy Institute says more than 5000 Haredi students were enrolled in rabbi-approved, gender-segregated, post–high school technological training programs during the 2021–22 school year, with an additional 15,600 students pursuing a college degree.

Although graduates of Haredi high schools have a weak foundation in math, science, and English, their rigorous religious education, which often involves extensive memorization and analysis, "has taught them how to learn," says educational neuroscientist Orit Elgavi-Hershler of the Mofet Institute, a teacher training consortium. "And they know how to think," she says, which can enable them to close educational gaps.

"On the ground, I see that the young generation of Haredi men and women want to be part of the future of the economy, and there are a lot of companies open to recruiting more Haredim," says Moshe Friedman, a Haredi entrepreneur who co-founded KamaTech, a training program that collaborates with some of Israel's largest technology firms. "I'm very optimistic."

Demographer Sergio DellaPergola, a professor emeritus at the Hebrew University of Jerusalem, sees promise in such initiatives. "If the Haredim get a general education, they'll be a highly productive addition to Israeli society," he says. "Everyone will benefit." ■

Michele Chabin is a journalist in Israel.

Emergency medicine doctor Saleena Subaiya (left) developed severe postvaccine symptoms, with their partner Lawrence Purpura, an infectious disease doctor.

COVID-19

# Studies probe COVID-19 shots' link to rare symptoms

## Details emerge for uncommon cases of neurologic complications, blood pressure swings, and other side effects

*By* **Gretchen Vogel** *and* **Jennifer Couzin-Frankel**

COVID-19 vaccines have saved millions of lives, and the world is gearing up for a new round of boosters. But like all vaccines, those targeting the coronavirus can cause side effects in some people, including rare cases of abnormal blood clotting and heart inflammation. Another apparent complication, a debilitating suite of symptoms that resembles Long Covid, has been more elusive, its link to vaccination unclear and its diagnostic features ill-defined. But in recent months, what some call Long Vax has gained wider acceptance among doctors and scientists, and some are now working to better understand and treat its symptoms.

"You see one or two patients and you wonder if it's a coincidence," says Anne Louise Oaklander, a neurologist and researcher at Harvard Medical School. "But by the time you've seen 10, 20," she continues, trailing off. "Where there's smoke, there's fire."

Cases seem very rare—far less common than Long Covid after infection. Symptoms can include persistent headaches, severe fatigue, and abnormal heart rate and blood pressure. They appear hours, days, or weeks after vaccination and are difficult to study. But researchers and clinicians are increasingly finding some alignment with known medical conditions. One is small fiber neuropathy, a condition Oaklander studies, in which nerve damage can cause tingling or electric shock–like sensations, burning pain, and blood circulation problems. The second is a more nebulous syndrome, with symptoms sometimes triggered by small fiber neuropathy, called postural orthostatic tachycardia syndrome (POTS). It can involve muscle weakness, swings in heart rate and blood pressure, fatigue, and brain fog.

Patients with postvaccination symptoms may have features of one or both conditions, even if they don't meet the criteria for a diagnosis. Both are also common in patients with Long Covid, where they're often attributed to an immune overreaction.

Although more researchers are now taking Long Vax seriously, regulators in the United States and Europe say they have looked for, but have not found, a connection between COVID-19 vaccines and small fiber neuropathy or POTS. "We can't rule out rare cases," says Peter Marks, director of the U.S. Food and Drug Administration's Center for Biologics Evaluation and Research, which oversees vaccines. "If a provider has somebody in front of them, they may want to take seriously the concept [of] a vaccine side effect," he says. But Marks also worries about "the sensational headline" that could mislead the public, and he emphasizes that vaccine benefits far outweigh any risks.

Despite the uncertainties, German Minister of Health Karl Lauterbach acknowledged in March that though rare, Long Covid–like symptoms after vaccination are a real phenomenon. He said his ministry was working to organize funding for studies, although none has been announced so far.

Researchers studying these complications also worry about undermining trust in COVID-19 vaccines. Harlan Krumholz, a cardiologist at Yale University, says concern that the antivaccine movement would seize on any research findings made him hesitant at first to dive in. But about a year ago he and Yale immunologist Akiko Iwasaki began to welcome postvaccination patients into a new study called LISTEN that also includes Long Covid patients. Among other things, it aims to correlate symptoms with immune cell patterns in blood samples.

"I'm persuaded that there's something going on" with these side effects, Krumholz says. "It's my obligation, if I truly am a scientist, to have an open mind and learn if there's something that can be done."

***SCIENCE* FIRST** wrote about these health concerns in January 2022, describing efforts by scientists at the National Institutes of Health to study and treat affected individuals (28 January 2022, p. 364). A study including 23 people was posted as a preprint in May 2022 but never published. Following *Science*'s story, almost 200 people contacted the journal sharing their postvaccination symptoms.

Research has since proceeded slowly. This is "a challenging outcome to monitor," Tom Shimabukuro of the U.S. Centers for Disease Control and Prevention told a government advisory committee in January. Still, more than two dozen case studies have accumulated describing POTS or small fiber neuropathy following a COVID-19 shot, regardless of the vaccine manufacturer.

In 2021, Sujana Reddy, now an internal medicine resident physician at East Alabama Health, published a case study of postvaccine POTS in a 42-year-old man. She then heard from more than 250 people describing similar health concerns. Reddy began a study documenting POTS cases, not yet published, which now includes 55 people who developed symptoms 1 to 2 weeks after vaccination.

Another team scanned for postvaccine POTS across a specific population. Researchers led by cardiologists Alan Kwan and Susan Cheng at Cedars-Sinai Medical Center analyzed a health database

of almost 285,000 people in the Los Angeles area; all had received at least one COVID-19 shot. They found that within 90 days after a shot, the rate of POTS-related symptoms was about 33% higher than in the 3 months before; 2581 people were diagnosed with POTS-related symptoms after vaccination, compared with 1945 beforehand. However, the study found a bigger effect from COVID-19 itself: The rate of POTS symptoms in about 12,000 unvaccinated people after infection was 52% higher than beforehand. Although Kwan cautions against extrapolating these numbers to a wider population, he says the pattern is intriguing. "Our data show a relatively clear signal that there probably is an increase in POTS after vaccination and after infection," he says.

Others find the study compelling. "Even last year I was a little bit cautious" about the link between POTS and vaccination, says Tae Chung, a neuromuscular physiatrist who runs the POTS clinic at Johns Hopkins University. "I didn't have quantitative data to back it up, but now I feel like I do." Still, Chung stresses that this paper and other data also suggest COVID-19 vaccines protect against POTS and other Long Covid symptoms, and he remains a strong advocate for vaccination.

**AN IMMUNE OVERREACTION** to SARS-CoV-2 spike protein, which COVID-19 vaccines use to induce protective antibodies, is one possible cause of these symptoms. One theory is that after vaccination some people generate another round of antibodies targeting the first. Those antibodies could function somewhat like spike itself: Spike targets a cell surface protein called the angiotensin-converting enzyme 2 (ACE2) receptor, enabling the virus to enter cells. The rogue antibodies might also bind to ACE2, which helps regulate blood pressure and heart rate, says Bernhard Schieffer, a cardiologist at the University of Marburg. If those antibodies disrupt ACE2 signaling, that could cause the racing heart rates and blood pressure swings seen in POTS.

Small fiber neurons also have the ACE2 receptor on their surface, so in theory rogue antibodies could contribute to neuropathy. But Matthew Schelke, a neurologist at Columbia University who has treated small fiber neuropathy in both Long Covid and postvaccine patients, says pinning down a connection won't be easy. Even when unusual antibodies turn up in someone's blood, "it's extremely difficult to know if any of these are pathogenic or if they are just bystanders," he says. Other immune system components that fuel inflammation may also harm nerves, he notes.

Some people appear susceptible to complications after both infection and vaccination, a double vulnerability that can put them in an agonizing situation—eschewing more vaccine doses, often on the advice of their doctors, but also fearing the hazards of Long Covid. Reddy falls into this category. She caught the virus in early 2020 while caring for an infected patient and developed Long Covid, including POTS. Her symptoms dramatically worsened after her first vaccine dose.

An emergency medicine physician in New York City, Saleena Subaiya, experienced the same phenomenon in reverse. Within 24 hours of the second dose of COVID-19 vaccine in January 2021, they developed severe cognitive impairment, fatigue, and a loss of balance. Subaiya improved somewhat over the next year, but was forced to switch to a part-time research job. Then in December 2021, they were hit by Omicron and relapsed.

Long Vax has become a focus for cardiologist Bernhard Schieffer, shown here treating a patient with a different condition.

PHOTO: UNIVERSITY HOSPITAL OF GIESSEN AND MARBURG

Postvaccination illness is "a long, relentless disease," says Lawrence Purpura, an infectious disease specialist at Columbia University who is also Subaiya's partner—and who treats both Long Covid patients and those with chronic symptoms after vaccination.

**A DIAGNOSIS OF POTS** or small fiber neuropathy post-vaccination can guide treatment. In POTS, doctors focus on increasing salt and fluid intake to boost blood volume and maintain blood pressure. Beta blockers, which slow racing hearts, may also help.

Small fiber neuropathy is treated with various medications to manage symptoms, and for severe cases sometimes intravenous immunoglobulin (IVIG), an expensive and hard-to-access antibody mixture that can tamp down immune overreactions. Some case studies report that IVIG helped people with postvaccine small fiber neuropathy, at least temporarily.

A more radical approach is plasma exchange, which is sometimes used for autoimmune disorders. Here the patient's plasma—the liquid part of the blood containing antibodies and proteins—is separated from the blood cells and discarded. The blood cells are then returned to the patient along with a replacement liquid.

Plasma exchange helped a man who developed small fiber neuropathy following his second dose of a COVID-19 vaccine, Schelke and colleagues reported in October 2022 in *Muscle & Nerve*. "He responded very well," with months of improvement, Schelke says, but recently returned with worsening symptoms.

Schieffer, meanwhile, has developed an experimental treatment regime that he says showed promise in an unpublished study of eight Long Covid and eight postvaccine patients. It includes statins, which can blunt inflammation in the circulatory system; and angiotensin II type 1 receptor blockers, which can help block the ACE2 pathway activation that may be dysregulated in patients. Schieffer and colleagues are hoping to start a clinical trial of the therapies with 500 people who have either Long Covid or postvaccine symptoms.

Patient advocates are trying to spur new studies. In the next few months, the nonprofit REACT19 plans to distribute small grants, in the tens of thousands of dollars, to teams studying immunology, biomarkers, and other features of postvaccine illness.

Even modest support matters, Krumholz says, because "it's incumbent on us to produce preliminary data" to win over funders with deep pockets. His and Iwasaki's LISTEN project now has about 2000 participants, roughly 1000 with Long Covid and 750 with postvaccination symptoms, he says. (LISTEN also includes healthy controls.) He hopes to begin to release results later this summer.

As countries draft plans for a round of updated vaccines, some scientists theorize, reassuringly, that people whose immune systems accepted a previous shot without incident would be highly unlikely to experience immune dysfunction following a booster. But that doesn't change the urgent need to help those suffering now, Reddy says. "We need to figure out why it's happening to this subset of people. Why is it happening to them and not everyone else?" ■

FEATURES

# THE STAR WATCHER

## Amateur astronomer Koichi Itagaki is one of the most prolific supernova hunters of all time

*By* **Dennis Normile**

For Koichi Itagaki, it was just another evening of supernova hunting. After his usual 7 p.m. dinner with his wife, he drove to his private observatory in the hills above his home in Yamagata, Japan, 290 kilometers north of Tokyo. He set out treats for the only visitors he allows on his celestial patrols: a stray cat he calls Nora and the raccoon dogs that warily approach from the surrounding forest. He then took a seat within his "headquarters," a cozy hut equipped with a bed, minifridge, microwave—and a dozen monitors used to control seven telescopes at three locations across Japan.

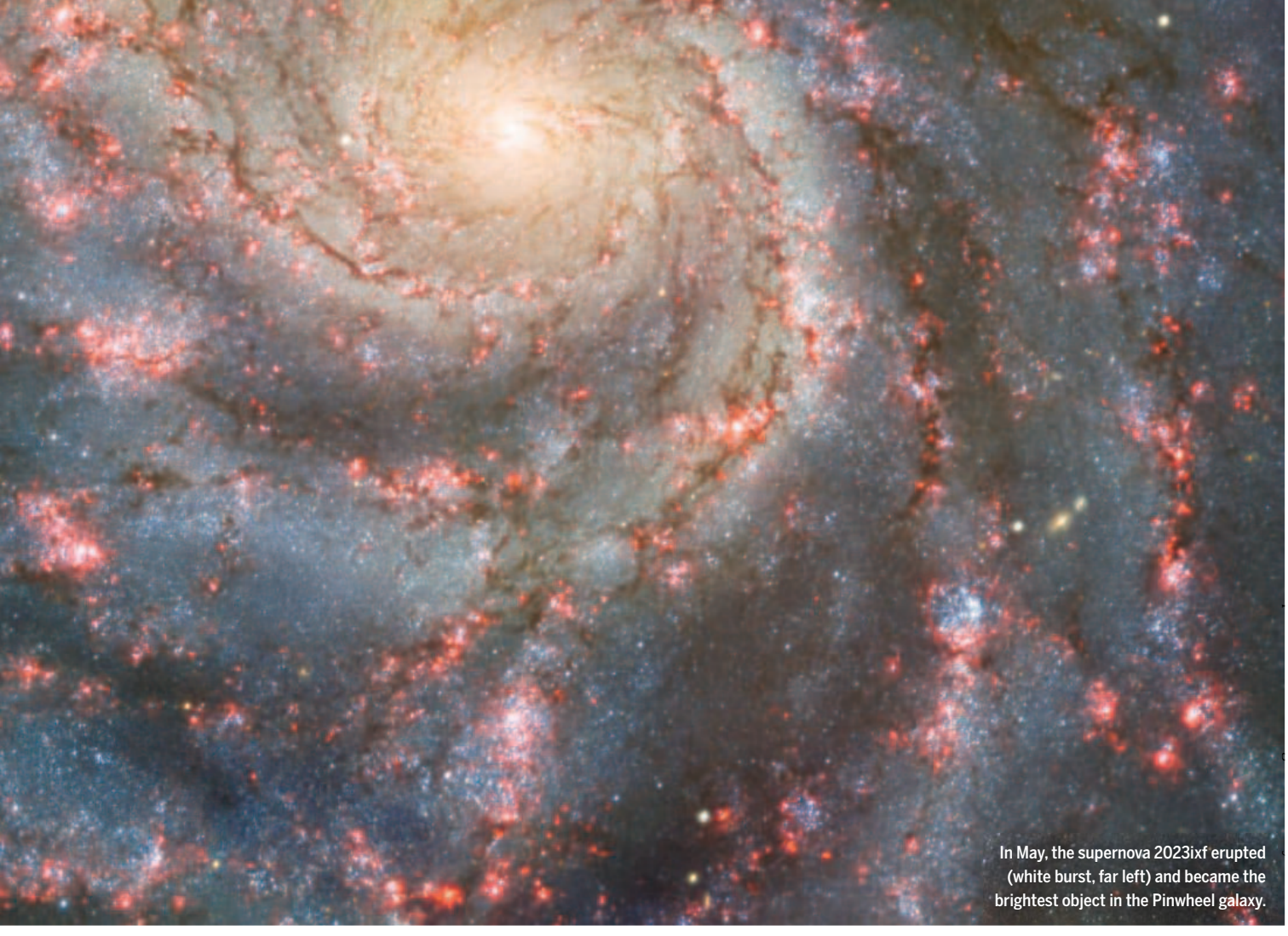
On clear nights, each telescope runs through a routine, focusing on one of the approximately 1000 galaxies Itagaki monitors for two 15-second exposures before swiveling to the next target. On this night, clouds blanketed most of the country, leaving a clear view only for his two telescopes in Okayama, 700 kilometers to the southwest. But as the clock ticked into the early morning of 20 May, clouds drifted over Okayama, too. Itagaki called it a night and drove home, leaving the telescopes on automatic in case the weather cleared.

It did. The next morning he had scanned the night's images for just 5 minutes when he spotted a new, bright object in a spiral arm of the Pinwheel galaxy, 21 million light-years away. "It was so bright, I thought there was no way this object could have been missed," he says. To his surprise, he was the first to post the news to the Transient Name Server (TNS), the International Astronomical Union's database of new celestial objects. As word spread on TNS and the Astronomer's Telegram, an email alert service, professionals and amateurs alike began pointing their instruments toward SN 2023ixf, the universe's newest exploding star and the closest to Earth in a decade. It was growing brighter by the minute.

For Itagaki, an amateur astronomer and semiretired snack food executive, such discoveries are routine. SN 2023ixf is his 172nd supernova, a total topped only by U.S.-based Tim Puckett, whose private observatory in Georgia has bagged at least 360 supernovae with the help of a worldwide network of volunteers who manually examine his images. Itagaki, by contrast, works alone. He "is one of the most prolific supernova observers in the world," says Andrew Howell, an astronomer at the University of California (UC), Santa Barbara.

But his achievement goes beyond numbers, Howell says. "The really special thing is that so many of the supernovae he finds turn out to be interesting." Although his formal education ended with high school, Itagaki has also co-authored a couple dozen scientific papers. "He's self-trained, so that's really incredible," says Daichi Hiramatsu, an astronomer at the Center for Astrophysics | Harvard & Smithsonian.

Amateurs have a long tradition of contributing to astronomy. Before the advent of large survey telescopes, they discovered many of the asteroids and comets that swing close to Earth. They have discovered galaxies and nebulae too dim and diffuse for large telescopes to notice. And they are now discovering planets outside the Solar System. But even among this storied population of hobbyists, Itagaki stands out "not only for his success, but also for his dedication," says Jeremy Shears, former president of the British Astronomical Association, which supports amateurs.



In May, the supernova 2023ixf erupted (white burst, far left) and became the brightest object in the Pinwheel galaxy.

**CHEERFUL, FRIENDLY**, and easy-going, Itagaki is a trim 75-year-old with wisps of white hair, wire-rimmed bifocals, and a self-effacing sense of humor. "I am not an astronomer," Itagaki says, smiling broadly while waving his left hand dismissively, as if shooing away a fly. "I'm looking for new celestial bodies as a hobby." He traces his path into astronomy to a boyhood fascination with lenses. "I used to play with lenses, using sunlight to burn paper," he says. In junior high, he spent his allowance on a DIY telescope kit and studied the Moon a bit. "I also used it to spy on the neighbors," he says, smiling and waving.

Then in 1963, a 19-year-old Japanese amateur named Kaoru Ikeya grabbed national headlines when he discovered a comet with a more substantial homemade telescope. "It amazed me that you could search the stars like this," Itagaki says. Within Itagaki's hut, a framed 1963 newspaper clipping of Ikeya's achievement hangs in homage.

After high school, Itagaki went to work at a local confectionary company, saving up to buy a 15-centimeter telescope with which he discovered his first comet at age 20. He joined Itagaki Peanuts, a snack food company started by his father, and eventually became its CEO. He is credited with being the first in Japan to sell snack "minipacks" that provide just a mouthful of peanuts, cashews, or candy-coated almonds. As the business prospered, Itagaki says he "invested a lot" in his passion—perhaps as much as the cost of typical Japanese house, he estimates. At age 60, he turned over operations of the 30-employee firm to his sons so he could devote himself to astronomy.

Like Ikeya and many other amateurs, Itagaki started out watching for comets. But in 1998, NASA set up a near-Earth object program, which enlisted major observatories to detect and track threatening celestial bodies. It left little room for amateur discoveries. So in 2000, "I gave up on comets and started searching for supernovae," Itagaki says.

To escape Yamagata's lights, Itagaki rented a plot in the hills above the city, and over time added the stainless-steel domes that now house his 60-, 50-, and 11-centimeter telescopes. About 8 years ago he set up the second observatory in Okayama, and 5 years ago he added a third observatory with another two telescopes in Kochi, 800 kilometers away on Shikoku, the smallest of Japan's four main islands.

The geographic distribution means at least one of the observatories is in the clear almost every night. "I don't think there's anyone else in the world who has a network of telescopes operated in such a way," says Hiramatsu, who adds that Itagaki also has an advantage in spotting supernovae that first appear during Japan's nighttime because Asia lacks a big survey. And whereas the big observatories tend to watch distant galaxies, Itagaki concentrates on nearby ones, where any supernovae can be studied when they are intrinsically fainter, says Azalee Bostroem, an astronomer at the University of Arizona. "This means they are discovered younger and can be observed later, sometimes years after the explosion," she says.

**SN 2023IXF SHOWS** how Itagaki gets the jump on other observers. Monika Soraisam, an astronomer at the Gemini Observatory, had written a program to automatically alert her to supernova candidates caught by the Zwicky Transient Facility, a survey using the 48-inch telescope at the Palomar Observatory in California. "I had the alert in my inbox the same night that Itagaki

Koichi Itagaki's homespun observatory in Yamagata is one of three he maintains across Japan in hopes of spotting supernovae before anyone else.

made his observations but did not have the time to follow up until later," she says.

Alex Filippenko, an astronomer at UC Berkeley, also missed the discovery. During a backyard star-viewing party the night of 18 May, he says he imaged the Pinwheel galaxy with an 11-centimeter telescope "just for fun." After hearing of SN 2023ixf's discovery, Filippenko found that he had serendipitously captured the supernova 11 hours before Itagaki. "Had I closely examined my images, I would have discovered it even earlier than Itagaki!" he says.

There is more at stake than bragging rights. "The sooner after the explosion that a supernova is detected, the more we can potentially learn about the exploding star and the explosion itself," Filippenko says. Four days after its discovery, SN 2023ixf peaked in brightness at about magnitude 11—still too faint to be seen with the naked eye but as bright as Proxima Centauri, the Sun's nearest neighbor. By then, numerous telescopes had tuned into the spectacle—including giants such as the 8.1-meter Gemini North telescope in Hawaii. NASA's Hubble and JWST space telescopes are also observing the supernova, which will fade over the course of many months.

Analysis of the light curve suggests the exploding star was a red supergiant hundreds of times bigger than the Sun. Having exhausted its nuclear fuel, the star's core collapsed, launching a shock wave that blew apart the star and produced a cosmic light show. In 2002 images from the Hubble telescope, Soraisam and her colleagues found what they believe is the progenitor. But she says they won't know for sure until "after the supernova has faded and we see that the star has vanished."

Howell says the critical data from Itagaki and other amateurs indicate an early, extra brightening of SN 2023ixf that supernova models don't predict. He points to three possible explanations for the excess light: a precursor explosion or eruption; dense shells of material around the star, lit up by the emerging shock wave; or unusual emissions from a nonspherical progenitor star. It is unlike anything seen before, "probably because we almost never catch supernovae this early," Howell says.

Itagaki's previous discoveries suggested it might be possible to see signs of unrest on a massive star even before it explodes. In 2004, he spotted a bright object in a spiral galaxy 77 million light-years away from Earth. No professionals verified his sighting before it disappeared 10 days later. Following a hunch it could be important, Itagaki periodically checked the location. In 2006, he (and two other amateur astronomers all working independently) discovered supernova 2006jc.

The explosion produced unexpected x-rays that Andrea Pastorello, an astronomer at Italy's Astronomical Observatory of Padova, thought were a clue to the 2004 outburst. Pastorello and his team concluded that what Itagaki saw was the progenitor star shedding its outer layers in a burst of light. Those layers lit up again with x-rays 2 years later when the supernova ejecta caught up and collided with the shed material.

It was previously thought that stars were quiet before going supernova. Itagaki's glimpse of that 2004 precursor event along with other observations is "really causing a big shift in our understanding of supernovae," Howell says.

**FOR DISCOVERING** and explaining the never-before-seen phenomenon, Pastorello's team and Itagaki were rewarded with a paper in *Nature* in 2007. It was Itagaki's first scientific publication, and supernova 2006jc remains his most memorable discovery. "But I had absolutely nothing to do with the paper; it just has my name on it," he says with that smile and self-deprecating wave.

Since then, Pastorello has worked with Itagaki on five other papers. "He is very collaborative, and his archive is a precious source of pre-explosion data," Pastorello says. That collaborative spirit has earned Itagaki co-authorship on what he estimates are "a lot" of papers, though he doesn't know how many. The Astrophysics Data System, operated by the Smithsonian Astrophysical Observatory, credits Itagaki as a co-author on 23 refereed publications and identifies scores more that cite his data. Papers from SN 2023ixf are likely to add to the tally.

SN 2023ixf has started its long fade out, but Itagaki's passion isn't dimming. Although already "superaged" by Japan's definition, he sees no reason to alter his routine. He's determined to catch a supernova going off in the Andromeda galaxy, the Milky Way's nearest neighbor, one that by cosmic averages is long overdue. He returned to his headquarters the night after finding SN 2023ixf for another solitary vigil.

"When I was in junior high school, I dreamed of building a hut, putting a big telescope in it, and having this life," he says. "That dream has come true." ■

PERSPECTIVES

MOLECULAR BIOLOGY

# An expanded view of transcription

## A new method expands chromatin to provide detailed images of transcription sites

By **Timothy J. Stasevich**[1] *and* **Hiroshi Kimura**[2]

How promoters and enhancers activate genes, and how they communicate to facilitate transcription, are unknown. The main transcriptional RNA polymerase, called RNA polymerase II (RNA Pol II), was originally thought to travel along a linear stretch of DNA. However, evidence gathered over the years paints a more complex picture, in which enhancer and promoter DNA recruit RNA Pol II and genes to larger superstructures, sometimes referred to as transcription factories, hubs, or condensates (*1*–*4*). The precise organization of these superstructures remains elusive, leaving many questions unanswered. On page 92 of this issue, Pownall *et al.* (*5*) describe a new form of expansion microscopy they call ChromExM, which they used to expand chromatin in zebrafish embryos to visualize transcription at high resolution. This expanded view provides evidence that promoters and enhancers come together during gene activation, only to be "kicked" apart soon after.

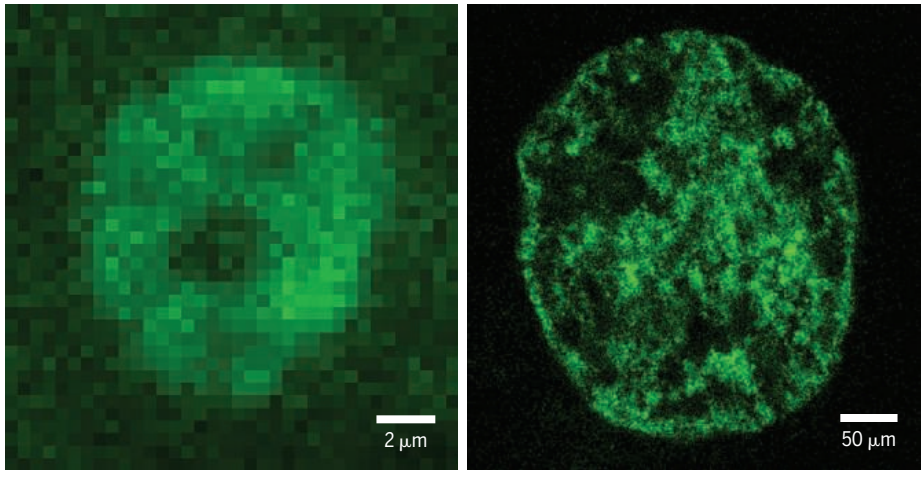
ChromExM modifies PanExM (*6*)—a technique to iteratively expand the full proteome using an entangled hydrogel mesh—to specifically expand chromatin 16-fold. For this, ChromExM uses a mesh that is optimized for catching and fluorescently labeling DNA-bound biomolecules such as histones, the major protein constituent of chromatin. Confocal imaging was used to visualize the labels, achieving a net spatial resolution of ~15 nm compared with the standard optical resolution of ~200 nm (see the figure). Pownall *et al.* also used stimulated emission depletion super-resolution microscopy to further zoom in on fluorescently labeled histones in the expanded zebrafish embryos, achieving a resolution of ~3 nm.

Chromatin forms a complex polymer, so it was unclear if chromatin architecture would be preserved after expansion. Although one previous study suggested that isotropic expansion is possible (*7*), other studies have demonstrated substantial artifacts after expansion (*8*). To test the ability of ChromExM to preserve chromatin architecture, Pownall *et al.* developed a method to optically etch equidistant parallel lines into chromatin. The lines remained essentially equidistant and parallel after expansion. They also confirmed that finer chromatin ultrastructure is preserved after expansion by comparing in vitro clusters of nucleosomes, which are basic units of chromatin consisting of DNA and histones, imaged with electron microscopy and ChromExM.

During the maternal–zygotic transition stages, when zygotic transcription is activated in a subset of genes, Pownall *et al.* observed that highly expressed microRNA gene clusters had hundreds of active RNA Pol II molecules arranged in a line. The authors also saw similar strings of RNA Pol II at nonspecific loci scattered across the nuclei. They interpreted each string as RNA Pol II molecules arranged in single file, generating nascent RNAs from a single gene. These strings are reminiscent of the RNA Pol II convoys that are predicted from models of transcription imaging in live cells (*9*). Previously, nascent RNAs have been observed arranged along string-like loops, but only in extremely long genes (*10*). Occasionally the strings observed by Pownall *et al.* were branched, suggesting multiple intertwined genes, indicative of a transcription factory or hub. In total, there were around 800 RNA Pol II strings per nucleus. Pownall *et al.* also observed ~30 larger "macroclusters" of RNA Pol II per nucleus, some connected to RNA Pol II strings, similar to the large and transient RNA Pol II clusters observed in living cells (*3, 11*).

Intriguingly, the macroclusters were sometimes filled with clusters of the transcription factor Nanog, but those clusters were physically distinct from the RNA Pol II clusters, suggesting that they bind to different sections of DNA. Pownall *et al.* hypothesized that those DNA elements were promoters and enhancers that were in close proximity before active RNA production by elongation. To test this hypothesis, they blocked elongation using an RNA Pol II inhibitor. This led to a significant reduction in the number of RNA Pol II strings, a closer association between Nanog and RNA Pol II, and an increase in RNA Pol II at promoters. These data indicate that there is an increase in transient enhancer–promoter contact when transcription factors cluster together but transcription is not activated (*12*).

This expanded view of transcription provides further support for the "kiss and run" model (*3*), where promoters and enhancers make transient contact to activate




**A fluorescent probe for DNA is used to visualize the nuclei from two different zebrafish embryos. The embryo on the right has undergone the ChromExM expansion process, providing a much more detailed view.**

[1]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO, USA. [2]Cell Biology Center, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, Japan. Email: tim.stasevich@colostate.edu; hkimura@bio.titech.ac.jp

genes. Furthermore, given that elongation inhibition led to an increase in promoter–enhancer association (kissing), Pownall *et al.* refine the model to "kiss and kick" to emphasize the active role elongation seems to be playing in separating promoters and enhancers. If elongation truly kicks apart enhancers and promoters, then the kick provides a simple explanation for why transcription would occur in well-defined bursts. The existence of this kick is supported by other observations. Phosphorylation of RNA Pol II at the onset of elongation can convert the association of RNA Pol II from transcription initiation factor condensates to splicing factor condensates, which are required for the processing of transcripts to mRNA (*4*). Indeed, transcription coactivators on enhancers and elongating RNA Pol II are physically separated (*13, 14*).

Looking ahead, it will be important to devise more direct tests of the kiss and kick model by imaging specific promoters and enhancers during ChromExM. This could be done by combining ChromExM with techniques for in situ labeling of endogenous DNA elements. This experiment would help establish whether the macroclusters observed by Pownall *et al.* correspond to a specific set of genes or if they represent a stage that all activated genes pass through.

Other questions should also be addressed. For example, can the RNA Pol II strings be seen in most genes without ChromExM, and what are their dynamics in live cells? It is also unknown how these findings relate to the three-dimensional organization of the genome, including chromatin domains and their stability, and whether these features are specific to developing embryos. In contrast to the early embryo—which has a rapid cell cycle and thus a robust transcription program that is executed in a short space of time—enhancer–promoter contact might not be so critical in somatic cells (*15*). Ultimately, it is exciting to see experiments beginning to converge on a unified model for transcriptional activation. ■


**REFERENCES AND NOTES**

1. F. J. Iborra, A. Pombo, D. A. Jackson, P. R. Cook, *J. Cell Sci.* **109**, 1427 (1996).
2. M. Mir *et al.*, *Genes Dev.* **31**, 1784 (2017).
3. W.-K. Cho *et al.*, *Science* **361**, 412 (2018).
4. Y. E. Guo *et al.*, *Nature* **572**, 543 (2019).
5. M. E. Pownall *et al.*, *Science* **381**, 92 (2023).
6. O. M'Saad, J. Bewersdorf, *Nat. Commun.* **11**, 3850 (2020).
7. E. L. Faulkner *et al.*, *J. Cell Sci.* **135**, jcs259009 (2022).
8. I. Kubalová *et al.*, *Chromosome Res.* **28**, 355 (2020).
9. K. Tantale *et al.*, *Nat. Commun.* **7**, 12248 (2016).
10. S. Leidescher *et al.*, *Nat. Cell Biol.* **24**, 327 (2022).
11. L. S. Forero-Quintero *et al.*, *Nat. Commun.* **12**, 3158 (2021).
12. K. Kuznetsova *et al.*, *Curr. Biol.* **33**, 164 (2023).
13. J. Li *et al.*, *Cell* **178**, 491 (2019).
14. H. Ohishi *et al.*, *Nat. Commun.* **13**, 7672 (2022).
15. J. P. Karr *et al.*, *Genes Dev.* **36**, 7 (2022).


10.1126/science.adi8187

PLANT BIOLOGY

# Maintaining asymmetry in cell division

## Promoting asymmetric division through microtubule dynamics establishes cell fate

*By* **Camila Goldy** *and* **Marie-Cécile Caillaud**

In multicellular organisms, cell division contributes to the generation of specialized cells. In walled organisms such as plants, the orientation of the division plane by which two daughter cells separate is of special importance because cells are embedded in a matrix and cannot relocate. To define the future cell division zone before mitosis, plants have evolved a specific structure called the preprophase band (PPB), which is composed of microtubules. In the context of asymmetric cell division, a polar domain is established at the plasma membrane to allow the differentiation of one of the daughter cells while maintaining a stem cell population. How plant cells establish the polar domain remains enigmatic. On page 54 of this issue, Muroyama *et al.* (*1*) reveal that a polarized destabilization of microtubules demarcates a region of the plant cell as incompatible with microtubule PPB establishment, thereby orienting asymmetric cell division and guiding the development of stomata (pores for gas exchange).

During leaf development, stomata are produced in the surface cell layer by iterative asymmetric cell divisions. This process involves the conversion of a subset of protodermal cells into meristemoid mother cells that divide asymmetrically to produce two daughter cells of unequal volumes. The small daughter cell becomes a meristemoid, which ultimately produces guard cells surrounding the stomatal pore while the larger daughter cell [the stomatal lineage ground cell (SLGC)] further divides asymmetrically. The daughter cells of the SLGC eventually differentiate into puzzle-shaped epidermal cells (pavement cells) while maintaining a controlled amount of meristemoids (*2*). During this process, a plasma membrane–associated polarity crescent defined by BREAKING OF ASYMMETRY IN THE STOMATAL LINEAGE (BASL) and BREVIS RADIX family (BRXf) proteins is formed and maintained throughout asymmetric cell division (*3*). The orientation of the subsequent asymmetric cell division, defined by the BASL-BRXf polarity crescent, is tightly controlled to enforce the "one-cell spacing rule," which ensures that stomata do not directly touch, thus optimizing gas exchange during photosynthesis while minimizing water loss (*4*).

To address the long-standing question of how the orientation of asymmetric divisions is determined and maintained, Muroyama *et al.* used stomatal lineage formation in *Arabidopsis thaliana* as a model. They showed that the cortical division zone (a landmark for the site of future division) and the microtubule PPB cannot form within the BASL-BRXf polarity crescent. This result agrees with earlier reports showing that the localization of BASL-BRXf is polarized before PPB formation when cortical division zone components are initially recruited to the cortex (*5–7*). Within asymmetrically dividing cells in the leaf epidermis, the default state for cell division orientation appears to be the shortest wall (requiring the least energy consumption). However, when the shortest path conflicts with the one-cell spacing rule, these rules need to be broken. The authors report that the polarized BASL-BRXf domain is required to override default division patterns during formative asymmetric divisions and that this is altered in *basl* mutants. Reciprocally, the establishment of the microtubule PPB is required for stomatal patterning. In a mutant that no longer forms the microtubule PPB, the division plane often bisects the BASL-BRXf polarity crescent, leading to aberrant cell fate in the resulting daughter cells.

How does the BASL-BRXf polar domain influence PPB positioning? Using time-lapse imaging in planta, Muroyama *et al.* demonstrate that the BASL-BRXf polar domain locally depletes cortical microtubules from the plasma membrane. There, the BASL-BRXf polar domain induces microtubule plus-end depolymerization rather than physically preventing microtubule polymerization into the domain per se. Based on these results, the authors conclude that polarization-mediated destabilization of the

Laboratoire Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRAe, Lyon, France. Email: marie-cecile.caillaud@ens-lyon.fr

genes. Furthermore, given that elongation inhibition led to an increase in promoter–enhancer association (kissing), Pownall *et al.* refine the model to "kiss and kick" to emphasize the active role elongation seems to be playing in separating promoters and enhancers. If elongation truly kicks apart enhancers and promoters, then the kick provides a simple explanation for why transcription would occur in well-defined bursts. The existence of this kick is supported by other observations. Phosphorylation of RNA Pol II at the onset of elongation can convert the association of RNA Pol II from transcription initiation factor condensates to splicing factor condensates, which are required for the processing of transcripts to mRNA (*4*). Indeed, transcription coactivators on enhancers and elongating RNA Pol II are physically separated (*13, 14*).

Looking ahead, it will be important to devise more direct tests of the kiss and kick model by imaging specific promoters and enhancers during ChromExM. This could be done by combining ChromExM with techniques for in situ labeling of endogenous DNA elements. This experiment would help establish whether the macroclusters observed by Pownall *et al.* correspond to a specific set of genes or if they represent a stage that all activated genes pass through.

Other questions should also be addressed. For example, can the RNA Pol II strings be seen in most genes without ChromExM, and what are their dynamics in live cells? It is also unknown how these findings relate to the three-dimensional organization of the genome, including chromatin domains and their stability, and whether these features are specific to developing embryos. In contrast to the early embryo—which has a rapid cell cycle and thus a robust transcription program that is executed in a short space of time—enhancer–promoter contact might not be so critical in somatic cells (*15*). Ultimately, it is exciting to see experiments beginning to converge on a unified model for transcriptional activation. ■

## REFERENCES AND NOTES


1. F. J. Iborra, A. Pombo, D. A. Jackson, P. R. Cook, *J. Cell Sci.* **109**, 1427 (1996).
2. M. Mir *et al.*, *Genes Dev.* **31**, 1784 (2017).
3. W.-K. Cho *et al.*, *Science* **361**, 412 (2018).
4. Y. E. Guo *et al.*, *Nature* **572**, 543 (2019).
5. M. E. Pownall *et al.*, *Science* **381**, 92 (2023).
6. O. M'Saad, J. Bewersdorf, *Nat. Commun.* **11**, 3850 (2020).
7. E. L. Faulkner *et al.*, *J. Cell Sci.* **135**, jcs259009 (2022).
8. I. Kubalová *et al.*, *Chromosome Res.* **28**, 355 (2020).
9. K. Tantale *et al.*, *Nat. Commun.* **7**, 12248 (2016).
10. S. Leidescher *et al.*, *Nat. Cell Biol.* **24**, 327 (2022).
11. L. S. Forero-Quintero *et al.*, *Nat. Commun.* **12**, 3158 (2021).
12. K. Kuznetsova *et al.*, *Curr. Biol.* **33**, 164 (2023).
13. J. Li *et al.*, *Cell* **178**, 491 (2019).
14. H. Ohishi *et al.*, *Nat. Commun.* **13**, 7672 (2022).
15. J. P. Karr *et al.*, *Genes Dev.* **36**, 7 (2022).


10.1126/science.adi8187

PLANT BIOLOGY

# Maintaining asymmetry in cell division

## Promoting asymmetric division through microtubule dynamics establishes cell fate

*By* **Camila Goldy** *and* **Marie-Cécile Caillaud**

In multicellular organisms, cell division contributes to the generation of specialized cells. In walled organisms such as plants, the orientation of the division plane by which two daughter cells separate is of special importance because cells are embedded in a matrix and cannot relocate. To define the future cell division zone before mitosis, plants have evolved a specific structure called the preprophase band (PPB), which is composed of microtubules. In the context of asymmetric cell division, a polar domain is established at the plasma membrane to allow the differentiation of one of the daughter cells while maintaining a stem cell population. How plant cells establish the polar domain remains enigmatic. On page 54 of this issue, Muroyama *et al.* (*1*) reveal that a polarized destabilization of microtubules demarcates a region of the plant cell as incompatible with microtubule PPB establishment, thereby orienting asymmetric cell division and guiding the development of stomata (pores for gas exchange).

During leaf development, stomata are produced in the surface cell layer by iterative asymmetric cell divisions. This process involves the conversion of a subset of protodermal cells into meristemoid mother cells that divide asymmetrically to produce two daughter cells of unequal volumes. The small daughter cell becomes a meristemoid, which ultimately produces guard cells surrounding the stomatal pore while the larger daughter cell [the stomatal lineage ground cell (SLGC)] further divides asymmetrically. The daughter cells of the SLGC eventually differentiate into puzzle-shaped epidermal cells (pavement cells) while maintaining a controlled amount of meristemoids (*2*). During this process, a plasma membrane–associated polarity crescent defined by BREAKING OF ASYMMETRY IN THE STOMATAL LINEAGE (BASL) and BREVIS RADIX family (BRXf) proteins is formed and maintained throughout asymmetric cell division (*3*). The orientation of the subsequent asymmetric cell division, defined by the BASL-BRXf polarity crescent, is tightly controlled to enforce the "one-cell spacing rule," which ensures that stomata do not directly touch, thus optimizing gas exchange during photosynthesis while minimizing water loss (*4*).

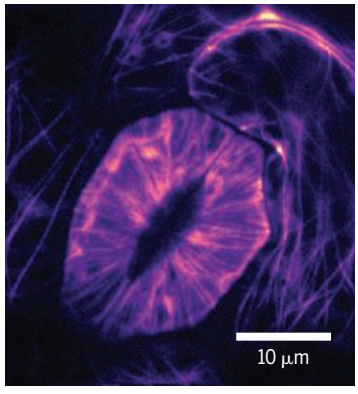
To address the long-standing question of how the orientation of asymmetric divisions is determined and maintained, Muroyama *et al.* used stomatal lineage formation in *Arabidopsis thaliana* as a model. They showed that the cortical division zone (a landmark for the site of future division) and the microtubule PPB cannot form within the BASL-BRXf polarity crescent. This result agrees with earlier reports showing that the localization of BASL-BRXf is polarized before PPB formation when cortical division zone components are initially recruited to the cortex (*5–7*). Within asymmetrically dividing cells in the leaf epidermis, the default state for cell division orientation appears to be the shortest wall (requiring the least energy consumption). However, when the shortest path conflicts with the one-cell spacing rule, these rules need to be broken. The authors report that the polarized BASL-BRXf domain is required to override default division patterns during formative asymmetric divisions and that this is altered in *basl* mutants. Reciprocally, the establishment of the microtubule PPB is required for stomatal patterning. In a mutant that no longer forms the microtubule PPB, the division plane often bisects the BASL-BRXf polarity crescent, leading to aberrant cell fate in the resulting daughter cells.

How does the BASL-BRXf polar domain influence PPB positioning? Using time-lapse imaging in planta, Muroyama *et al.* demonstrate that the BASL-BRXf polar domain locally depletes cortical microtubules from the plasma membrane. There, the BASL-BRXf polar domain induces microtubule plus-end depolymerization rather than physically preventing microtubule polymerization into the domain per se. Based on these results, the authors conclude that polarization-mediated destabilization of the

Laboratoire Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRAe, Lyon, France. Email: marie-cecile.caillaud@ens-lyon.fr

microtubule plus ends renders a region of the plasma membrane incompatible with microtubule PPB formation, thereby orienting cell division.

As in plant cells, division orientation in animal cells often prescribes daughter cell polarity by determining whether a cell inherits a cortical polar domain. For division-plane specification in animal cells, the orientation of the nucleus and/or mitotic spindle emerges from physical constraints that are sensed by microtubules, leading to bisection of the spindle axis in cytokinesis. To bias geometrical rules to break symmetry, polar domains typically recruit or activate cytoskeleton-associated proteins at the cortex, yielding net forces that reorient the mitotic structures such as the astral microtubules, which extend from the spindle poles (*8*). In the context of asymmetrical division in plant cells, the study by Muroyama *et al.* suggests that the orientation of the asymmetric division is not obtained through an effect of astral microtubules, because plants do not have them, but rather through effects on cortical microtubules transitioning toward the formation of a PPB before mitosis. Moreover, in plant asymmetric cell division, the polar domain does not specify a position for cytokinesis but acts as a forbidden landmark where microtubules cannot align, preventing the division plane from colliding with it.




**Super-resolved imaging shows guard cell cortical microtubules (pink) in stomata of *Nicotiana benthamiana*.**

The proposed mechanism by which polarity proteins could specify the division plane by creating a microtubule-depleted zone raises the questions of how the BASL-BRXf complex is recruited to the plasma membrane and how it excludes microtubules. Based on preliminary genetic evidence (*9*), BASL and BRXf are mutually required for the polarization of each other. In the current model, BRXf proteins reversibly associate with the plasma membrane through palmitoylation at their N termini (*9*). BASL is recruited through an interaction between BASL and the BRX domains that define the BRXf proteins. The nature of the upstream signal that defines the location of BASL-BRXf polarization remains to be discovered.

It is likely that the scaffold proteins BASL and BRXf recruit a microtubule-associated protein to affect microtubule dynamics rather than directly mediating plus-end depolymerization. The critical microtubule-associated proteins might not be cell division specific because the same microtubule depletion is observed after division in SLGCs for as long as the polar domain persists. But how can the component responsible for the direct effect on microtubules be identified? The establishment of asymmetry in animal cells often involves both biochemical and mechanical processes (*8*). Because microtubules were proposed to be tension sensors (*10*), the mechanical constraints and subsequent signaling might be aspects to consider.

During pavement cell differentiation, specific small guanosine triphosphatases (GTPases) called Rho of Plants (ROPs) act as master regulators of cell polarity, mediating the recruitment of microtubule-associated proteins in the polar domain to shape cytoskeleton distribution. Upstream, ROPs are regulated by several stimuli, including the receptor-like kinase FERONIA (FER) pathway, which participates in sensing and/or transducing mechanical signals (*11*–*13*). Investigating the cross-talk between ROP signaling and symmetry breaking in the context of asymmetric cell division might be key in the future. Overall, the study by Muroyama *et al.* highlights the singular mechanism used by plants to control the inheritance of fate regulators by connecting polarity with microtubule interactions. ■


## REFERENCES AND NOTES

1. A. Muroyama, Y. Gong, K. S. Hartman, D. C. Bergmann, *Science* **381**, 54 (2023).
2. X. Guo, L. Wang, J. Dong, *New Phytol.* **232**, 60 (2021).
3. L. R. Lee, D. C. Bergmann, *J. Cell Sci.* **132**, jcs228551 (2019).
4. L. J. Pillitteri, J. Dong, *Arabidopsis Book* **11**, e0162 (2013).
5. A. Muroyama, Y. Gong, D. C. Bergmann, *Curr. Biol.* **30**, 4467 (2020).
6. Y. Gong *et al.*, *Development* **148**, dev199919 (2021).
7. X. Guo, C. H. Park, Z.-Y. Wang, B. E. Nickels, J. Dong, *Nat. Plants* **7**, 667 (2021).
8. N. Ierushalmi, K. Keren, *Curr. Opin. Cell Biol.* **72**, 91 (2021).
9. M. H. Rowe, J. Dong, A. K. Weimer, D. C. Bergmann, *bioRxiv* 10.1101/614636 (2019).
10. O. Hamant, D. Inoue, D. Bouchez, J. Dumais, E. Mjolsness, *Nat. Commun.* **10**, 2360 (2019).
11. C. Zhang *et al.*, *Curr. Biol.* **32**, 518 (2022).
12. W. Tang *et al.*, *Curr. Biol.* **32**, 508 (2022).
13. T. Lauster *et al.*, *Curr. Biol.* **32**, 532 (2022).



## ACKNOWLEDGMENTS

The authors thank Y. Jaillais and G. Ingram for their input on the manuscript. M.-C.C. is supported by French National Research Agency grant ANR-20-CE13-0026-02. C.G. is funded by a European Molecular Biology Organization (EMBO) fellowship (ATLF 646-2021).

10.1126/science.adi6664


PHYSICS

# Probing fundamental particles with molecules

## Molecular spectroscopy constrains the size of the electron's electric dipole moment

*By* **Mingyu Fan** *and* **Andrew Jayich**

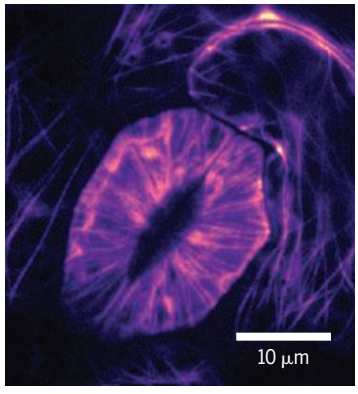
The discovery of the Higgs boson in 2012 marked the completion of the Standard Model (SM) of particle physics. The SM remains the best theory of the known fundamental particles and how they interact, but there is evidence that the SM is incomplete. A shortcoming is that the SM predicts nearly equal amounts of matter and antimatter after the Big Bang, which would result in a cosmos full of light after the matter and antimatter particles mutually annihilate. The physicist Andrei Sakharov outlined the basic conditions that are needed to create the matter imbalance, called the baryon asymmetry of the Universe (BAU), that allows for our existence (*1*). One of the Sakharov conditions, time symmetry violation (TSV), suggests that other particles and interactions beyond those of the SM await discovery. On page 46 of this issue, Roussy *et al.* (*2*) report the most stringent limits to date on the size of the electron's permanent electric dipole moment (EDM), which may help explain the BAU.

Time symmetry ensures that when we imagine reversing time, the physical laws follow their time-reversed paths. TSV is needed to generate the BAU because it is a necessary condition to convert antimatter to matter, which may have happened in the early Universe just after the Big Bang, leaving more matter than antimatter. After annihilation of particles with their antiparticles, the excess amount of matter formed the current Universe. The physics that describes the SM mostly obeys time sym-

Department of Physics, University of California, Santa Barbara, CA, USA. Email: mingyufan212@gmail.com; jayich@gmail.com

microtubule plus ends renders a region of the plasma membrane incompatible with microtubule PPB formation, thereby orienting cell division.

As in plant cells, division orientation in animal cells often prescribes daughter cell polarity by determining whether a cell inherits a cortical polar domain. For division-plane specification in animal cells, the orientation of the nucleus and/or mitotic spindle emerges from physical constraints that are sensed by microtubules, leading to bisection of the spindle axis in cytokinesis. To bias geometrical rules to break symmetry, polar domains typically recruit or activate cytoskeleton-associated proteins at the cortex, yielding net forces that reorient the mitotic structures such as the astral microtubules, which extend from the spindle poles (*8*). In the context of asymmetrical division in plant cells, the study by Muroyama *et al.* suggests that the orientation of the asymmetric division is not obtained through an effect of astral microtubules, because plants do not have them, but rather through effects on cortical microtubules transitioning toward the formation of a PPB before mitosis. Moreover, in plant asymmetric cell division, the polar domain does not specify a position for cytokinesis but acts as a forbidden landmark where microtubules cannot align, preventing the division plane from colliding with it.




**Super-resolved imaging shows guard cell cortical microtubules (pink) in stomata of *Nicotiana benthamiana*.**

The proposed mechanism by which polarity proteins could specify the division plane by creating a microtubule-depleted zone raises the questions of how the BASL-BRXf complex is recruited to the plasma membrane and how it excludes microtubules. Based on preliminary genetic evidence (*9*), BASL and BRXf are mutually required for the polarization of each other. In the current model, BRXf proteins reversibly associate with the plasma membrane through palmitoylation at their N termini (*9*). BASL is recruited through an interaction between BASL and the BRX domains that define the BRXf proteins. The nature of the upstream signal that defines the location of BASL-BRXf polarization remains to be discovered.

It is likely that the scaffold proteins BASL and BRXf recruit a microtubule-associated protein to affect microtubule dynamics rather than directly mediating plus-end depolymerization. The critical microtubule-associated proteins might not be cell division specific because the same microtubule depletion is observed after division in SLGCs for as long as the polar domain persists. But how can the component responsible for the direct effect on microtubules be identified? The establishment of asymmetry in animal cells often involves both biochemical and mechanical processes (*8*). Because microtubules were proposed to be tension sensors (*10*), the mechanical constraints and subsequent signaling might be aspects to consider.

During pavement cell differentiation, specific small guanosine triphosphatases (GTPases) called Rho of Plants (ROPs) act as master regulators of cell polarity, mediating the recruitment of microtubule-associated proteins in the polar domain to shape cytoskeleton distribution. Upstream, ROPs are regulated by several stimuli, including the receptor-like kinase FERONIA (FER) pathway, which participates in sensing and/or transducing mechanical signals (*11*–*13*). Investigating the cross-talk between ROP signaling and symmetry breaking in the context of asymmetric cell division might be key in the future. Overall, the study by Muroyama *et al.* highlights the singular mechanism used by plants to control the inheritance of fate regulators by connecting polarity with microtubule interactions. ■


## REFERENCES AND NOTES

1. A. Muroyama, Y. Gong, K. S. Hartman, D. C. Bergmann, *Science* **381**, 54 (2023).
2. X. Guo, L. Wang, J. Dong, *New Phytol.* **232**, 60 (2021).
3. L. R. Lee, D. C. Bergmann, *J. Cell Sci.* **132**, jcs228551 (2019).
4. L. J. Pillitteri, J. Dong, *Arabidopsis Book* **11**, e0162 (2013).
5. A. Muroyama, Y. Gong, D. C. Bergmann, *Curr. Biol.* **30**, 4467 (2020).
6. Y. Gong *et al.*, *Development* **148**, dev199919 (2021).
7. X. Guo, C. H. Park, Z.-Y. Wang, B. E. Nickels, J. Dong, *Nat. Plants* **7**, 667 (2021).
8. N. Ierushalmi, K. Keren, *Curr. Opin. Cell Biol.* **72**, 91 (2021).
9. M. H. Rowe, J. Dong, A. K. Weimer, D. C. Bergmann, *bioRxiv* 10.1101/614636 (2019).
10. O. Hamant, D. Inoue, D. Bouchez, J. Dumais, E. Mjolsness, *Nat. Commun.* **10**, 2360 (2019).
11. C. Zhang *et al.*, *Curr. Biol.* **32**, 518 (2022).
12. W. Tang *et al.*, *Curr. Biol.* **32**, 508 (2022).
13. T. Lauster *et al.*, *Curr. Biol.* **32**, 532 (2022).



## ACKNOWLEDGMENTS

The authors thank Y. Jaillais and G. Ingram for their input on the manuscript. M.-C.C. is supported by French National Research Agency grant ANR-20-CE13-0026-02. C.G. is funded by a European Molecular Biology Organization (EMBO) fellowship (ATLF 646-2021).

10.1126/science.adi6664


PHYSICS

# Probing fundamental particles with molecules

## Molecular spectroscopy constrains the size of the electron's electric dipole moment

*By* **Mingyu Fan** *and* **Andrew Jayich**

The discovery of the Higgs boson in 2012 marked the completion of the Standard Model (SM) of particle physics. The SM remains the best theory of the known fundamental particles and how they interact, but there is evidence that the SM is incomplete. A shortcoming is that the SM predicts nearly equal amounts of matter and antimatter after the Big Bang, which would result in a cosmos full of light after the matter and antimatter particles mutually annihilate. The physicist Andrei Sakharov outlined the basic conditions that are needed to create the matter imbalance, called the baryon asymmetry of the Universe (BAU), that allows for our existence (*1*). One of the Sakharov conditions, time symmetry violation (TSV), suggests that other particles and interactions beyond those of the SM await discovery. On page 46 of this issue, Roussy *et al.* (*2*) report the most stringent limits to date on the size of the electron's permanent electric dipole moment (EDM), which may help explain the BAU.

Time symmetry ensures that when we imagine reversing time, the physical laws follow their time-reversed paths. TSV is needed to generate the BAU because it is a necessary condition to convert antimatter to matter, which may have happened in the early Universe just after the Big Bang, leaving more matter than antimatter. After annihilation of particles with their antiparticles, the excess amount of matter formed the current Universe. The physics that describes the SM mostly obeys time sym-

Department of Physics, University of California, Santa Barbara, CA, USA. Email: mingyufan212@gmail.com; jayich@gmail.com

metry; although there is TSV in the SM, it is not nearly enough to account for the BAU. TSV interactions arising from undiscovered particles can generate permanent EDMs of SM particles such as the electron. An EDM in an electric field causes an energy shift when its orientation is flipped with respect to the field. An electron EDM (eEDM) violates time-reversal symmetry: If time were reversed, its magnetic moment would flip and the EDM would not, looking fundamentally different from before time-reversal (see the figure). All eEDM measurements so far are consistent with zero, indicating that no particle has been found to induce enough TSV to generate an eEDM that is within the current experimental sensitivities. However, tighter experimental bounds on the eEDM probes increasingly higher particle masses, analogous to how higher-energy particle colliders can search for more-massive particles.

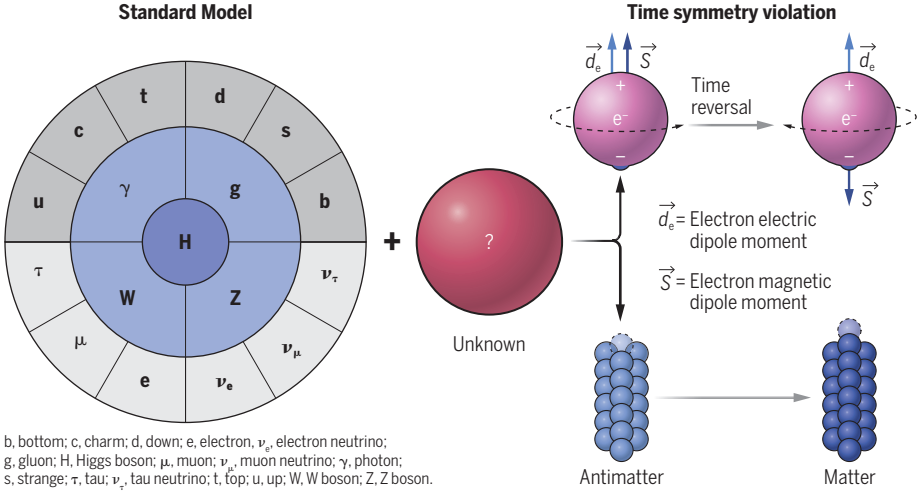
Roussy *et al.* measured eEDM by precision spectroscopy of $HfF^+$ molecules. They produced and trapped $HfF^+$ molecules in an ion trap and polarized them with a rotating electric field. The polarization controls the orientation of the large internal molecular field that the molecule's valence electron experiences. With an additional applied magnetic field, lasers, microwaves, and radiofrequency fields, Roussy *et al.* measured the energy of $HfF^+$ molecules in two configurations: when the electron's magnetic moment was aligned and when it was anti-aligned to the molecule's massive electric field. They probed the molecular energy shift that arises between the two configurations, allowing them to limit the eEDM to be no larger than $4.1 \times 10^{-30}$ $e$ cm. With their vacuum apparatus that fits on a table, they were able to probe for undiscovered particles up to a mass of 40 TeV, which is 10 times the energy of particles than can be observed in the Large Hadron Collider at CERN, which costs about $4.75 billion to build and $1 billion to run annually. Considerable effort by Roussy *et al.* went into meticulously studying their experimental apparatus and measurement technique so they could understand systematic uncertainties in minute detail to ensure that no spurious signals were mistakenly introduced (*3*).

Sensitivity in the detection of the eEDM increased when molecules were used instead of atoms (*4*). The two previous eEDM limits come from brief measurements of many thorium oxide (ThO) molecules (*5*, *6*) using a cryogenic beam as compared to the trapped molecular ion approach of Roussy *et al.*, in which fewer molecules are measured for much longer times. The combined ThO and $HfF^+$ measurements set more stringent limits on sources of TSV than either measurement does on its own. This is because measurements in different molecules have different sensitivities to underlying sources of TSV. In the event of a nonzero eEDM measurement, a combined analysis of multiple measurements in different systems narrows down the contributions from the individual TSV sources and provides more stringent bounds to test theories. Knowledge from EDM measurements across multiple systems would help guide the requirements of a future high-energy particle collider that could create the TSV particles responsible for the BAU.

The SM predicts a very small eEDM of less than $10^{-38}$ $e$ cm, which is roughly a billion times smaller than the limit reported by Roussy *et al.* This extremely small SM contribution to the eEDM gives future experiments operating at higher sensitivities plenty of room for discovery. In parallel with the ThO and $HfF^+$ eEDM experiments, there has been progress toward next-generation TSV measurements that aim to improve sensitivity by using more-sensitive molecular species, larger numbers of a given species, and longer measurement coherence times. These include using polyatomic molecules (*7*), exotic isotopes with deformed nuclei (*8*), and even probing TSV in solids where larger numbers of atoms are trapped in a crystal (*9*). Technological improvements to allow such advances include laser cooling and trapping of polyatomic molecules (*10*), spectroscopy of radioactive molecules (*11*), and fine-grained control of molecular ions with quantum logic spectroscopy (*12*, *13*). A combination of such approaches could allow the detection of high-energy particles perhaps at PeV energies (1 PeV = $10^{15}$ eV), and along the way the sought-for undiscovered particles may be detected. Such experiments will constrain high-energy theories beyond the SM. But upgrades of both the ThO beam experiment and the trapped molecular ion technique, including a switch from $HfF^+$ to $ThF^+$, are poised to enhance the particle-mass reach of eEDM experiments in the near future. ■

## Trying to understand the missing matter

The Standard Model does not explain important problems such as the baryon asymmetry of the Universe (BAU), which requires undiscovered particles for additional time symmetry violation (TSV) interactions to convert antimatter to matter in the early Universe. Such particles may be discovered by precision measurements of the electron electric dipole moment, which is a TSV moment induced by an asymmetry of the charge distribution in the electron.




b, bottom; c, charm; d, down; e, electron, $\nu_e$, electron neutrino; g, gluon; H, Higgs boson; μ, muon; $\nu_\mu$, muon neutrino; γ, photon; s, strange; τ, tau; $\nu_\tau$, tau neutrino; t, top; u, up; W, W boson; Z, Z boson.

GRAPHIC: A. FISHER/*SCIENCE*


### REFERENCES AND NOTES

1. A. D. Sakharov, *JETP Lett.* **5**, 24 (1967).
2. T. S. Roussy *et al.*, *Science* **381**, 46 (2023).
3. L. Caldwell *et al.*, arXiv:2212.11837 (2023).
4. J. J. Hudson *et al.*, *Nature* **473**, 493 (2011).
5. The ACME Collaboration, *Science* **343**, 269 (2014).
6. V. Andreev *et al.*, *Nature* **562**, 355 (2018).
7. I. Kozyryev, N. R. Hutzler, *Phys. Rev. Lett.* **119**, 133002 (2017).
8. J. T. Singh, *Hyperfine Interact.* **240**, 29 (2019).
9. H. D. Ramachandran, A. C. Vutha, arXiv:2304.10331 (2023).
10. N. B. Vilas *et al.*, *Nature* **606**, 70 (2022).
11. R. F. Garcia Ruiz *et al.*, *Nature* **581**, 396 (2020).
12. C. W. Chou *et al.*, *Nature* **545**, 203 (2017).
13. C. W. Chou *et al.*, *Science* **367**, 1458 (2020).



### ACKNOWLEDGMENTS

The authors thank S. Thomas for helpful discussions and the US Department of Energy (DE-SC0022034) and NSF (PHY-2146555) for support.

10.1126/science.adi8499

SOLAR CELLS

# Tandems have the power

## Perovskite-silicon tandem solar cells break the 30% efficiency threshold

A laboratory-scale perovskite-silicon tandem solar is shown; the blue aperture defines the device area.

*By* **Stefaan De Wolf** *and* **Erkan Aydin**

One way to increase the efficiency of a solar cell is to optimize it to the wide spectrum of sunlight for conversion into electricity. In this vein, a tandem solar cell may comprise a top cell that absorbs high-energy photons and a bottom cell that absorbs low-energy photons. Such a device can be fabricated as a stack of layers, with the top and bottom cell interconnected to each other. Monolithic integration of a perovskite solar cell and silicon solar cell into a tandem device is a promising path toward high-performance photovoltaics (PVs) at affordable cost. On pages 59 and 63 of this issue, Chin *et al.* (*1*) and Mariotti *et al.* (*2*), respectively, report different approaches to achieving power conversion efficiencies (PCEs) in perovskite-silicon tandem solar cells of >30%. Overcoming this threshold provides confidence that high-performance, low-cost PVs can be brought to the market.

In 2022, the cumulative installed global PV capacity reached 1.2 TW (*3*). Yet to avert the catastrophic scenarios associated with global warming, the total PV capacity needs to increase to ~75 TW by 2050 (*4*). Over the past decades, growth of the mainstream PV market, which is dominated by silicon technologies, has largely been fueled by a systematic increase in the performance of PV cells and modules (cells connected together in strings and packaged to form larger units), combined with enhanced module reliability and scaling for manufacturing. Collectively, these factors have driven down the levelized cost of electricity (costs over the system's lifetime divided by its energy production) for PVs. Modern commercial silicon solar cells now reach PCEs >24% (*5*), and the best laboratory cell has a PCE of 26.8% (*6*). This is attributed to innovations such as surface and contact passivation—that is, a means to suppress the recombination of charge carriers (electrons and holes). However, with a maximum theoretical PCE of 29.4% (*7*), practical performance limits for a single-junction silicon PV (a solar cell made from a silicon wafer as light absorber) are in sight. To sustain further decreases in levelized cost of electricity of commercial PVs, scalable higher-performance solar cell concepts with reliability that is equivalent to silicon products need to come to maturity quickly.

Tandem solar cells are the most straightforward route toward decreasing the levelized cost of electricity, well beyond what is possible for single-junction solar cells. This is because tandem cells can drastically lower charge-carrier thermalization. Thermalization is an energy exchange process of carriers that results in heat dissipation, which limits the efficiency of a solar cell. As well, optical transmission loss limits the performance of single-junction solar cells with a given bandgap energy, $E_G$. Thermalization losses are associated with incident photons with energy > $E_G$; transmission losses are associated with photon energies < $E_G$. Tandem cells minimize these losses by stacking solar cells, positioning the cell with the larger bandgap toward the sunward side. With a bandgap of 1.12 eV and a proven track record as an affordable, mature PV technology, silicon is well suited to act as a bottom cell in a tandem device. The ideal top cell bandgap for a two-junction tandem is in the 1.65 to 1.75 eV range.

Metal halide perovskites have emerged in the past 15 years as a promising semiconductor class for multiple optoelectronic applications, especially solar cells. Because their bandgap can be tuned to cover the range desirable for top cells, their integration with silicon into tandems has quickly become an active area of research. Toward this end, several approaches for depositing perovskites onto silicon have been developed that are based on solution processes, vacuum deposition, or a combination of both methods.

Initial work on monolithic tandem cells focused on the development of low-temperature deposition processes for the top cell to guarantee compatibility with the high-efficiency bottom cell. The latter is often based on silicon heterojunction technology. Tandem cells are now designed to collect electrons on their sunward side and holes on the rear side because this maximizes electric current flow. Such tandem polarity makes use of specific contact layers with minimized loss of current due to light absorption without charge generation (known as parasitic optical absorption) (*8*). To further enhance the current density, bottom cells with textured surfaces (both sunward and rear sides) have been increas-

King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah, Saudi Arabia.
Email: stefaan.dewolf@kaust.edu.sa

ingly adopted to enhance light trapping (*9*). Nevertheless, to fully realize the potential of perovskite-silicon tandems, irrespective of how the perovskite is deposited, a main hurdle has been the minimization of charge loss from recombination of electrons and holes in the perovskite top cell, in particular at its interface with the electron-collecting contact stack at the light-facing side. This contact stack usually features a $C_{60}$ fullerene layer (an electron transport layer) that efficiently extracts electrons from the perovskite layer. In previous works, insertion of a thin inorganic film between the perovskite and $C_{60}$ layer proved beneficial to minimize voltage losses (*8*, *10*) but left residual defects on the perovskite surface unpassivated.

Both Chin *et al.* and Mariotti *et al.* report different strategies that resolve the issue of top contact passivation through molecular approaches, resulting in certified >30% efficient perovskite-silicon tandems (see the figure). This efficiency threshold has been broken for different types of perovskite-silicon tandems, pointing to the broad design space of perovskite technology.

In 2022, Chin *et al.* announced the first monolithic perovskite-silicon tandem to exceed the 30% efficiency barrier, with a certified PCE of 31.25% (*11*). They now describe how they achieved this with a hybrid two-step perovskite deposition method that combines thermal evaporation of an inorganic template followed by a solution-based conversion of this scaffold into the perovskite phase. The conversion is accomplished through the intercalation of organic constituents. This occurs on the textured sunward side of the silicon bottom cell that is textured on both sides. However, because of the high film conformality of this process, the top surface of the perovskite remains textured too. As a result, spin-casting (using centrifugal force) a solution onto this surface as a passivation approach is not directly applicable to this device structure because it would lead to inhomogeneous molecular coverage on the textured top surface. This would adversely affect device performance. Chin *et al.* overcame this challenge by blending a phosphonic acid–based additive [FBPAc (2,3,4,5,6-pentafluorobenzylphosphonic acid)] to the precursor solution in the spin-casting step. The bulky FBPAc molecules are then expelled toward the top surface during perovskite crystallization, yielding larger crystal domains. Moreover, once expelled, FBPAc molecules passivated surface defects, reducing charge loss from recombination at the perovskite-$C_{60}$ top interface. This substantially improved tandem performance.

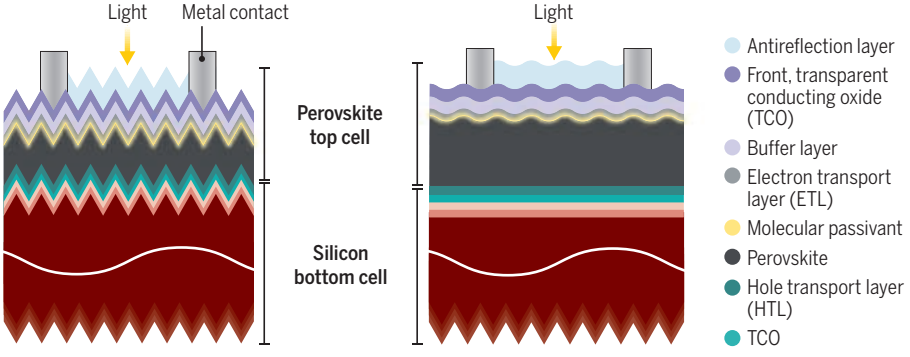
Following a different approach, Mariotti *et al.* show that the perovskite-$C_{60}$ interface can also be passivated through insertion of piperazinium iodide, an ionic liquid, for tandems that use a solution-processed perovskite on a front-flat silicon bottom cell. Piperazinium iodide was chosen because it contains electron acceptor and electron donor functional groups that can interact with positive and negative perovskite surface defects, respectively. This modification improved the energetic band alignment, reduced recombination losses, and enhanced electron extraction at the perovskite-$C_{60}$ interface. The resulting perovskite-silicon tandem demonstrated impressive open-circuit voltages of up to 2.00 V. The authors thus provide an explanation for their 2022 announcement of a perovskite-silicon tandem with certified PCE of 32.5%.

As next steps, it remains important to further improve the performance of perovskite-silicon tandems that could achieve a PCE well above 35%. At the same time, focus is needed on improving device reliability. Arguably, the most critical factor lies in the annual degradation rate under actual outdoor conditions, which for perovskite-silicon tandems remains largely unknown. To be commercially viable, this degradation rate should be on par with mainstream PV technologies, which is less than 1% relative per year. Furthermore, scaling from the current 1 $cm^2$ record-device areas to >100 $cm^2$, as is common in mainstream silicon manufacturing, is also critical. Notably, a certified record PCE of 28.6% for a commercial-sized device was recently announced by Oxford PV (*12*). To be market competitive, in a high-throughput environment (manufacturing thousands of devices per hour), scaled tandems likely should show a net PCE benefit of at least 3% absolute on average, compared with the best commercial silicon cells (which is a continuously moving target).

Reaching these performance, stability, and scaling goals in about 5 years—which is the time it may take for commercial silicon PV to reach its practical PCE limit—is ambitious. But full engagement toward such ambitions would not only expedite the maturity of an exciting renewable energy technology but, more importantly, accelerate achieving the critical 75-TW global PV capacity target. ■

## Tandem solar cell structures

A tandem solar cell consists of a silicon bottom cell and a perovskite top cell. Perovskite absorbs blue light (high-energy photons) best, whereas silicon absorbs red light (low-energy photons). This combination maximizes the capture and conversion of sunlight into electricity more efficiently than that by single-cell types (single-junction solar cell). Contact layers are required to extract current from the tandem and need to be engineered toward maximal light-incoupling and minimal charge losses. Two new fabrication methods achieve cells with a power conversion efficiency of >30%. Both position a molecular passivant at the perovskite-electron transport layer interface of the front contact, minimizing charge recombination.




**Hybrid-processed on textured**
A hybrid, two-step process deposits perovskite conformally onto the silicon cell. The device is textured on both the sunward and rear sides. The perovskite mimics the surface texturing.

**Solution-processed on planar**
A spin-casting solution process deposits onto a largely smooth silicon surface. The textured rear enhances light trapping of red light.

GRAPHIC: N. CARY/*SCIENCE*

### REFERENCES AND NOTES


1. X. Y. Chin *et al.*, *Science* **381**, 59 (2023).
2. S. Mariotti *et al.*, *Science* **381**, 63 (2023).
3. https://iea-pvps.org/wp-content/uploads/2023/04/IEA_PVPS_Snapshot_2023.pdf.
4. N. M. Haegel *et al.*, *Science* **380**, 39 (2023).
5. Y. Chen *et al.*, *Prog. Photovolt. Res. Appl.* 10.1002/pip.3626(2022).
6. H. Lin *et al.*, *Nat. Energy* 10.1038/s41560-023-01255-2 (2022).
7. A. Richter, M. Hermle, S. W. Glunz, *IEEE J. Photovolt.* **3**, 1184 (2013).
8. K. Bush *et al.*, *Nat. Energy* **2**, 17009 (2017).
9. F. Sahli *et al.*, *Nat. Mater.* **17**, 820 (2018).
10. J. Liu *et al.*, *Science* **377**, 302 (2022).
11. https://www.nrel.gov/pv/cell-efficiency.html.
12. https://www.oxfordpv.com/news/oxford-pv-sets-new-solar-cell-world-record.


10.1126/science.adi6278

POLICY FORUM

ANIMAL WELFARE

# Animal welfare: Methods to improve policy and practice

## New methods are emerging to quantify human and animal welfare on a common scale, creating new tools for policy

*By* **Mark Budolfson[1,2,3], Bob Fischer[4,5], Noah Scovronick[6]**

There is growing international consensus that animal welfare is a crucial consideration in policy analysis, affecting domains ranging from food systems to biomedical research. Concern for animal welfare also features in many government regulations, certification programs, and institutional ethics codes across the globe and is central to many philanthropic and values-based investment decisions. However, although there are well-developed quantitative tools for incorporating human welfare into policy analysis, comparable tools for animal welfare are in their earliest stages. Without them, it is impossible to assess the net welfare impacts of a policy on humans and nonhumans alike on a common scale, which is crucial for making informed and transparent trade-offs (*1*). In practice, then, animal welfare is often ignored. Given that animal welfare matters in many cases, there is an urgent need for best-practice methods for integrating animal welfare into decision analyses.

The goal of this policy forum is threefold: first, to identify some important policy questions involving animal welfare that remain unanswered; second, to crystalize the limitations of existing tools for making progress on these questions, owing to the difficulty of directly comparing human and animal welfare on a common scale (as well as comparisons of animals of different species on a common scale); and third, to outline new methods that are emerging to overcome these challenges. We hope this analysis will accelerate refinement of these methods, increase interest in their deployment in policy analysis, and encourage broad interdisciplinary collaboration toward that end.

[1]Department of Environmental and Occupational Health and Justice, Rutgers University, New Brunswick, NJ, USA. [2]Center for Population-Level Bioethics, Rutgers University, New Brunswick, NJ, USA. [3]Population Wellbeing Initiative, University of Texas at Austin, Austin, TX, USA. [4]Department of Philosophy, Texas State University, San Marcos, TX, USA. [5]Rethink Priorities, San Francisco, CA, USA. [6]Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA. Email: mark.budolfson@rutgers.edu

## IMPORTANT POLICY QUESTIONS

Several policy questions illustrate where animal welfare is of crucial importance but where the absence of a common scale for assessing welfare inhibits progress.

### Health and biomedicine

Governmental agencies sometimes cull wild animals to reduce zoonotic disease transmission [e.g., Australia's culling of bats to reduce the threat of Hendra virus (*2*)]. When, if ever, does harm to wildlife provide sufficient reason to accept a higher level of human risk (*3*)? How should frameworks such as One Health and Planetary Health—which aim to improve the health of people, animals, and their shared environment—balance trade-offs within and across species?

Animal models are standard in biomedical research, though the rates at which therapies that work in animals translate to humans are often poor (*4*). Given that existing regulations call on researchers to reduce the number of animals used, when possible, it is important to compare animal welfare costs with the expected value of research. How can this be done rigorously?

### Domesticated animals, food, and climate change

No-kill animal shelters are popular because of public discomfort with euthanizing unwanted pets. Should animal shelters be publicly subsidized to avoid having to euthanize pets—subsidies that would benefit animals but would divert resources from other human interests?

Sustainable intensification in agriculture involves improving farm productivity without increasing inputs or land use. Unfortunately, many forms of sustainable intensification have negative implications for animal welfare, often due to high stocking densities (*5*). When, if ever, is it be[illegible] to increase environmental sustainability if it reduces animal welfare?

Estimates of the social cost of carbon, which quantify damages from climate change in economic terms, have improved considerably over time as a wider range of harms have been incorporated (*6*). However, the welfare impacts from climate change on both wild and domesticated animals have largely been ignored, partly because of the difficulty of quantifying those impacts (*7*).

### Trade and taxation

The General Agreement on Tariffs and Trade (GATT) protects free trade. Some countries have pushed for exceptions to block the import of low-welfare animal products (*8*). Is it appropriate to reduce economic efficiency for the sake of advancing animal welfare?

There have been calls for taxes on low-welfare animal products (*9*). However, without methods for quantifying animal welfare, it is unclear how to price negative welfare impacts from case to case.

### Conservation biology and resource management

New Zealand's Predator Free NZ initiative, which aims to eliminate all non-native predators from the country, has been hotly debated because of some of the wildlife control methods on which it relies. When, if ever, do the welfare impacts of such population management methods, such as anticoagulant poisons, outweigh the benefits of promoting the restoration of native ecosystems?

During the COVID-19 lockdowns, many urban animals starved because of the interruption of normal human activities. What level of risk, if any, should humans accept to provide assistance in such instances where there may be conflicts between the promotion of human and animal welfare? Which animals should be prioritized insofar as the aim is to minimize suffering?

### Philanthropic investment and prosocial decision-making

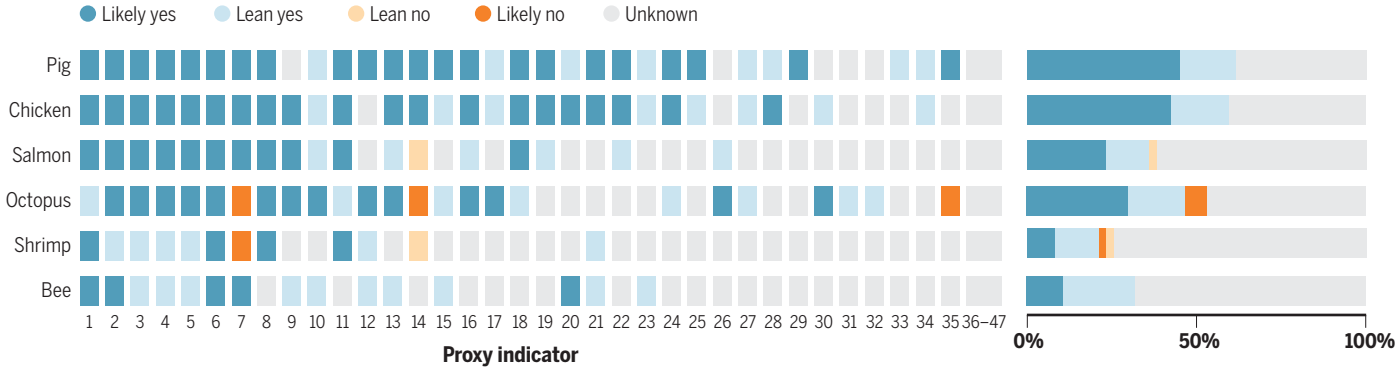
Given the prosocial goal of improving welfare as much as possible, how should individuals, businesses with corporate social responsibility programs, and major charitable organizations divide resources between animal welfare and humanitarian causes such as the United Nations Sustainable Development Goals?

## LIMITS OF EXISTING TOOLS

These issues raise several challenges, two of which have been the main focus of animal welfare science over the past several decades. First, animal welfare science has tried to identify conditions that negatively affect the

## Mapping animal welfare proxies

The figure outlines 47 potential animal welfare proxies for which assessments can be scored, weighted, and aggregated to produce relative estimates of species' welfare potential. Colored segments indicate whether the capacity is likely present based on existing evidence (*12*) (left). A simple method would be to give a binary (yes or no) score to each proxy and weight them all equally; species with a larger number of "yes" scores would imply a greater relative welfare potential in cross-species comparisons (right). More sophisticated methods of scoring, weighting, and aggregation are needed to reflect specific theories of welfare and to address uncertainties.




The proxies are as follows: 1– Taste-aversion behavior, 2– Reward-based learning, 3– Anxiety-like behavior, 4– Fear-like behavior, 5– Flexible self-protective behavior, 6– Response modified by painkillers, 7– Parental care, 8– Exhaustion-like behavior, 9– Associative learning from pain, 10– Self-medication, 11– Prioritizes pain response in relevant context, 12– Relief learning, 13– Play behavior, 14– Maternal response to offspring distress, 15– Valuing behavior, 16– Effects on exploratory behavior from pain, 17– Hyperalgesia, 18– Panic-like behavior, 19– Depression-like behavior, 20– Social buffering, 21– Displacement behavior, 22– Loneliness-like behavior, 23– Helping behavior, 24– Boredom-like behavior, 25– Emotional contagion, 26– Trace conditioning and pain, 27– Emotional reactions to learning, 28– Disgust-like behavior, 29– Friendship-like behavior, 30– Sensory-affective dissociation, 31– Curiosity-like behavior, 32– Joy-like behavior, 33– Posttraumatic stress disorder–like behavior, 34– Rescue behavior, 35– Play vocalization, 36– Concept of death, 37– Embarrassment-like behavior, 38– Envy-like behavior, 39– Guilt-like behavior, 40– Jealousy-like behavior, 41– Liking-wanting dissociation, 42– Love-like behavior, 43– Mourning-like behavior, 44– Pride-like behavior, 45– Sadness-like behavior, 46– Shame-like behavior, and 47– Sympathy-like behavior.

welfare of individual animals, such as crowding of caged or penned animals or restrictions that prevent the freedom to engage in species-specific natural behaviors. Second, it has tried to identify strategies for mitigating these negative effects (e.g., providing environmental enrichment for farmed animals). These goals have resulted in the generation of considerable new knowledge about the needs of animals and how to improve their welfare, as shown by influential frameworks such as the Five Domains model and the Welfare Quality model, which are representative of a rich literature that identifies factors relevant to animals' quality of life.

However, these frameworks cannot answer the policy questions described above. First, they are silent on the relative importance of different types of welfare impacts on individual animals even within a single species, preventing overall judgments about the relative goodness or badness of policy options. For example, rearing chickens bred to grow more slowly lowers the incidence of keel bone fractures but also increases the amount of time that chickens spend in intensive confinement: How should we evaluate the trade-off between the duration and intensity of suffering? Likewise, a move from caged to cage-free systems for chickens improves freedom of movement but results in a higher mortality rate due to higher incidences of smothering and injurious pecking. When is this trade-off acceptable?

Second, even if these frameworks were extended to address these intraspecies challenges, they would still be silent on how to compare the overall welfare impacts on members of one species with the overall welfare impacts on members of another species, including humans. This issue is a crucial impediment to answering questions of the type listed above, given that increased benefits to farmed or wild animals are often only possible at some cost to humans or to each other.

What is needed, therefore, is a method for aggregating, on a single scale, the net effects of a policy on the welfare of both animals and humans alike (as well as animals of different species). This would enable principled intraspecies and interspecies welfare comparisons, which are the key challenges for next-generation animal welfare methods.

### NEXT-GENERATION METHODS

Researchers are beginning to develop next-generation methods for quantifying and comparing both animal and human welfare on a single scale [e.g., (*7*, *10*–*12*)]. Work on these methods is expanding rapidly, fueled by increasing recognition of their importance for answering policy questions such as those listed above. As a result, the time is ripe for researchers across disciplines to join the effort to develop and refine these methods. Useful collaboration might include not only refinement of methods to quantify animal welfare but also careful formulation of policy questions where animal welfare is relevant, as well as clear articulation of the needs of policy analysts for incorporating animal welfare estimates.

There are several strategies for making progress. Some methods focus on anthropocentric valuations of animal welfare, such as consumer willingness to pay for given welfare improvements (*13*, *14*). Others try to provide intrinsic valuations of animal welfare by combining measurable empirical indicators that may be correlated with welfare potential across species (*10*, *12*), where welfare potential is a measure of how good and bad the life of an individual of a given species can be. For example, given the differences between them, including differences in their respective cognitive capacities and degrees of sociality, a pig can likely suffer more than a python (*12*).

Estimates of welfare potential might be based on hypothesized neural or other proxies for conscious states, cognitive and behavioral traits, or more complex multidimensional proxies (see the figure). For instance, some researchers have proposed defining a function directly linking neuron counts in the forebrain to welfare potential (*10*), resulting in a single score per species that can be used to make interspecies welfare comparisons. Alternatively, other research groups have proposed multidimensional proxies that can also be aggregated to arrive at a single overall welfare potential score (*12*). Analogous methods are used in economics to define functions from income (or other proxies) to individual human welfare. Though they are imperfect and must be used judiciously, keeping broader moral, legal, and political considerations in mind, such proxies can be valuable tools in some decision contexts.

Consider, for instance, two policies that reduce greenhouse gas emissions from animal agriculture by the same amount. Both focus on intensification, but one targets pork and the other targets chicken. With estimates of the welfare potential of each species, assessments of how their quality of life will change under each policy, and the number of ani-

GRAPHIC: D. AN-PHAM/*SCIENCE* BASED ON B. FISCHER (*12*)

mals affected, it becomes possible to estimate which policy is better for animal welfare.

Valuations that can inform policy may combine the results of several methods because there are many uncertainties to navigate. There are a number of dimensions about which there may be disagreements over the extent to which those dimensions contribute to welfare (see the figure). Analogous empirical and normative challenges have been addressed in other policy domains, such as what discount rate to use for analyses that require long time horizons (e.g., in climate policy) or how to value human mortality and morbidity. Sometimes addressing these challenges requires methods such as sensitivity analysis—where policies are evaluated across a range of normative and empirical views—that allow progress while acknowledging the existence of important disagreements. In the discounting debate, for instance, sensitivity analyses test how the optimal climate policy might differ given different assumptions about how much additional dollars of avoided climate damages generate well-being for the poor compared with the rich. Analogously, sensitivity tests reveal how the optimal policy regarding agricultural practices might differ given higher welfare potential estimates for animals compared with lower estimates. For example, some approaches estimate high welfare potential for chickens (*12*), whereas others estimate a much lower value (*10*). Some policies may be robustly recommended under either approach, such as low-cost interventions that greatly improve quality of life in intensive chicken farming operations. Other policies may not be as robust across approaches, such as relatively expensive interventions that have a smaller impact on quality of life. Still, given the importance of animal welfare, some policy guidance is preferable to none at all. Although research continues to uncover the intelligence, sociality, and affective capacities of animals, there is evidence to suggest that important gradients in such capacities exist across species. Moreover, as researchers devote more concerted attention to the challenges associated with interspecies trade-offs, progress can be made on how to provide actionable guidance to policy analysts, decision-makers, and stakeholders.

## ACCELERATING APPLICATION

With the preceding discussion in mind, we offer the following recommendations.

### Research across disciplines to develop and refine methods

Methods for making quantitative interspecies welfare estimates are in their infancy. Welfare economists, empirically minded philosophers, and animal welfare scientists have begun researching these methods, but this work could be substantially enriched by deep collaborations among diverse researchers, including comparative psychologists, neuroscientists, population biologists, and many others.

### Prioritize readily measurable empirical correlates of welfare

To enable interspecies comparisons, these correlates need to be specific enough that they represent genuine proxies for welfare but general enough that they can be applied to a wide range of species, thus allowing for sufficiently comprehensive analyses. Examples of correlates include the extent to which responses to noxious stimuli are modified by painkillers, which provides evidence for the ability to feel pain (see the figure).

### Develop flexible tools that are sensitive to normative disagreements

Policies that improve overall welfare must generally make some trade-offs. Decision analysis tools are often designed to rank policies as better or worse, such as many of the integrated assessment models that evaluate climate, air quality, or food systems policies. Accordingly, they make normative assumptions that can be controversial, such as those described above about discount rates or how to value mortality or morbidity. The same issue will arise when researchers incorporate animal welfare into such decision analysis tools or when developing new tools focused on animal welfare. Given important disagreements about the relative importance of human and animal welfare and the relative importance of different animals' welfare, researchers need to develop transparent tools that allow stakeholders to determine whether results are robust across reasonable ethical positions in the decision context. It is crucial to develop methods for incorporating animal welfare directly into such decision analyses so that they can serve their intended functions (i.e., assessing policies as better and worse).

### Discuss animal welfare in mainstream policy analysis

Animal welfare has long been neglected in policy choices, despite its ethical importance and the public's concern for it. However, it will be harder to ignore in the future because so many actions have impacts on animals and attitudes are steadily shifting in proanimal directions (*15*). As policy debates continue, it is essential to highlight impacts on animals as one among many relevant considerations. ■

Emerging methods can be used to estimate welfare impacts on livestock of high stocking density.


#### REFERENCES AND NOTES

1. J. Sebo, *Saving Animals, Saving Ourselves* (Oxford Univ. Press, 2022).
2. A. Linder, D. Jamieson, *Science* **379**, 621 (2023).
3. N. Sekar, D. Shiller, *Science* **369**, 629 (2020).
4. L. Kramer, R. Greek, *Bus. Soc. Rev.* **123**, 3 (2018).
5. T. Garnett *et al*., *Science* **341**, 33 (2013).
6. D. Diaz, F. Moore, *Nat. Clim. Change* **7**, 774 (2017).
7. K. Kuruc, J. McFadden, *Soc. Choice Welfare* 10.1007/s00355-023-01451-9 (2023).
8. K. Sykes, *World Trade Rev*. **13**, 471 (2014).
9. G. Perino, H. Schwickert, *Nat. Food* **4**, 160 (2023).
10. M. Budolfson, D. Spears, in *The Routledge Handbook of Animal Ethics*, B. Fischer, Ed. (Routledge, 2019), pp. 92–101.
11. R. Espinosa, N. Treich, *Soc. Choice Welfare* **56**, 531 (2021).
12. B. Fischer, The Welfare Range Table (2022); https://rethinkpriorities.org/publications/the-welfare-range-table.
13. C. J. Lagerkvist, S. Hess, *Eur. Rev. Agric. Econ*. **38**, 55 (2011).
14. F. B. Norwood, J. L. Lusk, *Compassion, by the Pound* (Oxford Univ. Press, 2011).
15. D. C. Fulton, M. J. Manfredo, J. Lipscomb, *Hum. Dimens. Wildl*. **1**, 24 (1996).



#### ACKNOWLEDGMENTS

M.B. and N.S. received support from Longview Philanthropy. B.F. received support from Open Philanthropy and Rethink Priorities. We thank J. Elsey for his work on the figure. All data and materials used are freely available online in (*12*).

10.1126/science.adi0121

Human food sources, such as dumpsters, increase the potential for human–bear conflict.

PHOTO: RGB VENTURES/SUPERSTOCK/ALAMY STOCK PHOTO

ANIMAL BIOLOGY

# The future of bears

## A deep dive into the eight living bear species emphasizes humanity's role in their fate

*By* **Özgün Emre Can**

On a cold but bright day in the winter of 2015, I found myself near the border between Nepal and Tibet questioning my decision to embark on a field study in one of the most remote places on Earth. The challenges and risks inherent in the work of a conservation biologist are not trivial. Yet as I contemplated how I had come to find myself in this position, I noticed three children playing next door with an unexpected object: a fluffy teddy bear. I snapped a photograph that I would later use to remind myself and my team of the importance of our efforts to build a better future for people and for wildlife.

Three years after completing that Himalayan field study, I felt similar surprise and joy while reading *Eight Bears*, which tells the stories of the planet's eight extant bear species and includes a delightful diversion into the origin of the teddy bear. The book, written by journalist Gloria Dickie—"a child obsessed with wild creatures" and an alumna of the University of Colorado Boulder—is the product of a long journey the author took around the world to see the planet's eight bear species in their natural habitats and to interview researchers who study them. The book's chapters are organized according to Dickie's travel route, which began in South America, where she learned about spectacled bears, more commonly known as Andean bears (*Tremarctos ornatus*), in Ecuador and Peru; then moved to Asia, where her investigations turned to sloth bears (*Melursus ursinus*) in India, giant pandas (*Ailuropoda melanoleuca*) in China, and sun bears (*Helarctos malayanus*) and moon bears, or more accurately Asiatic black bears (*Ursus thibetanus*), in Vietnam; and then to North America, where she explored the lives of American black bears (*Ursus americanus*) and brown bears (*Ursus arctos*) in the United States and polar bears (*Ursus maritimus*) in Canada.

**Eight Bears: Mythic Past and Imperiled Future**
*Gloria Dickie*
Norton, 2023. 272 pp.

This well-researched book is a pleasure to read and is packed with facts many readers will be surprised to learn. For example, in Britain, the royals once had a "Master of the Bears" to manage activities related to brown bears in London, and bears were used in bearbaiting—an activity that involved forcing bears to fight with a pack of aggressive dogs—until the British Parliament banned the practice in 1835. However, bear dancing—in which a tame bear is coerced to perform for entertainment purposes—was allowed to continue in Britain until 1911. Meanwhile, Dickie recounts how, in America, the explorer William Clark wrote in his diary about "a verry large and a turrible looking animal" in 1805, and how dozens of brown bears were killed during Clark's famous expedition with Meriwether Lewis. The book also introduces readers to the mythology that has long surrounded bears, including a folktale in which an Andean bear assumes the disguise of a human man in order to kidnap young women. Through these and other stories, the author succeeds in portraying the diversity of relationships people in different cultures have established with bears over thousands of years.

Dickie also has a good understanding of the problems occurring between bears and the people who coexist with them. In North America, she reveals, the availability of anthropogenic food sources is the main driver of conflicts with American black bears and brown bears. In Asia, meanwhile, expansion of people into bear habitats and habitat loss are major drivers of conflicts with Andean bears and sun bears. Dickie describes how sloth bears sometimes cause injuries and fatalities, resulting in retaliatory killings. As she rightly concludes, human–bear conflict, whether perceived or real, is shaping the future of bears.

Dickie's detailed, accurate, and holistic view of human–bear relationships is impressive. Even bear researchers can fall short in this regard, focusing on narrow scientific questions with little consideration of the big picture. The book's greatest strength thus lies in how the author taps into the wisdom of multiple, internationally recognized researchers and how she digests and aggregates their collective insights to give readers a comprehensive understanding of the lives of modern bears. The only things missing, in my opinion, were a world map showing the range of all eight bear species and the locations of the field sites the author visited; representative photographs of each species; and photographs of people who coexist with bears around the globe.

In *Eight Bears*, Dickie succeeds in meeting her goal of presenting "an unromanticized view of bears' tenuous position in the Anthropocene." "Losing bears," she notes, "would mean we lose a beautiful and complex relationship that has paralleled our own journey in this world." ■

The reviewer is at the Department of Biology, Ankara University, Beşevler, Ankara 06560 , Turkey. Email: oecan@ankara.edu.tr

10.1126/science.adi4084

Shane Campbell-Staton examines "The Gift of Maize" in Tlaxcala, Mexico.

ANTHROPOCENE

# Altering Earth in our image

## A new documentary series traces humanity's profound effects on the world around us

*By* **Mary Ellen Hannibal**

Shane Campbell-Staton is an evolutionary biologist at Princeton University, but you might mistake him for a seasoned journalist in *Human Footprint*, a six-part series set to air weekly on PBS beginning on 5 July 2023. The series, hosted by Campbell-Staton, reveals how, in our relatively brief history, humankind has wrought profound transformations to ecosystems, to other species, and even to our own DNA.

*Human Footprint* begins with an episode dedicated to species that have been introduced into ecosystems where they did not evolve. Viewers watch as Campbell-Staton gamely rides along with snake hunter Donna Kalil in the Everglades, on the hunt for nonnative pythons released into the wild via the Florida pet trade. The prodigious python eats prey as large as deer and is busily denuding this delicate wetland of its historic denizens.

Kalil likens pythons to aliens. They regenerate lost teeth and can regrow hearts and stomachs, the latter of which they sometimes fill with an alligator. "Once gators are on the menu," Campbell-Staton observes, "you've reached official gangster status."

How exactly to deter invasives is also a question asked about feral horses. Wild horses are descended from ancestors bred by humans for speed and quick reproductive turnover. Modern horses can double their population every 4 to 5 years, and proliferation of these beautiful animals is devastating the American West.

Although state and federal agencies use lethal deterrence to cull harmful wildlife populations, they are restricted by law from controlling horses this way. The reasoning is not rational. As author David Philipps puts it, horses are woven "very deeply into our idea of who we are ourselves," and of course we do not sanction shooting people either. Campbell-Staton reflects that "our desire to keep wild horses in the West might be rooted in our culture, but its outcome is biological." In the Anthropocene, this profound truth is evident everywhere.

In a nicely calibrated segue, Campbell-Staton takes on predation in the next episode. "Our ancestors were hunters before they were even humans," he tells viewers, remarking that pursuing prey has shaped our bodies, brains, and culture.

Our decimation of whale populations in pursuit of oil for fuel and food, he reveals, resulted in ecosystem depletion that continues to reverberate in today's oceans. "When whales are not there," says biologist Nick Pyenson, "they are not fertilizing the ocean." Food webs are decimated, and the ocean's chemistry is altered.

At Mozambique's Gorongosa National Park, Campbell-Staton looks for elephants with ecologist Dominique Gonçalves. Gonçalves runs a program to help elephant populations rebound after a devastating civil war reduced their numbers from 2500 to 200 in 16 years. Here, Campbell-Staton employs his evolutionary biologist chops to help explain how intensive hunting by humans changed elephant genetics. Half the females born to survivors of this elephanticide were tuskless, he reveals, three times as many as before the war, when elephant tusks were targeted for their high monetary value. [He has published peer-reviewed science documenting this phenomenon (*1*).]

Is humanity an overwhelming force of destruction? Not if you are a dog. Episode 3 details the evolution of wolf descendants to "Man's Best Friend." Biologist Brian Hare hypothesizes that wolves adapted themselves to humans, rather than the other way around. Today, that looks like an excellent decision. There are ~250,000 wild wolves on Earth at present, many of them persecuted by humans. Their descendants that took another genetic path now number around 1 billion, and many of them are living large indeed.

Episodes 4 and 5 delve further into the intimate intertwining of humanity with species that we have wittingly or unwittingly taken along on a wild ride of human domination. Here, viewers learn that livestock domesticated by humans for our consumption outweigh all wild mammals and birds combined by a factor of 10. In one sense, these species won the lottery of survival, but chickens, for example, probably would not consider the efficient ways we breed, raise, and slaughter them much of a prize.

The problem of how to feed so many of us without subjugating wild species is not addressed, but Campbell-Staton does cover a nifty effort to restrain domestic cats, who kill an estimated 1 to 3 billion birds and 7 to 10 million small mammals per year. These creatures, "the worst invasive species alive," might one day be deterred by AI algorithms in collars that recognize when a cat is about to hunt and emit a bird alarm.

**Human Footprint**
*Day's Edge Productions*
PBS, Wednesdays, 5 July to 9 August 2023, 9 p.m. EST (check local listings)

The final episode of *Human Footprint* considers cotton, and here Campbell-Staton departs somewhat from the narrative approach he has taken thus far. The fate of more than 5 million West Africans stolen from their homeland to farm cotton in the American South had more than economic and ecological impacts. Campbell-Staton visits a geneticist from 23andMe, where massive genetic datasets show that more than a significant percentage of contemporary African Americans have white, male ancestry. The specter of such widespread and systematic rape adds a new dimension to the series' repeated mantra: "What we do reveals who we truly are." ■

The reviewer is the author of *Citizen Scientist: Searching for Heroes and Hope in an Age of Extinction* (The Experiment, 2016). Email: maryellenhannibal@gmail.com

**REFERENCES AND NOTES**

1. S. C. Campbell-Staton *et al.*, *Science* **374**, 483 (2021).

10.1126/science.adi8075

# INSIGHTS

LETTERS

ILLUSTRATION: HANNAH AGOSTA

# Succinct discoveries that inspire

We asked young scientists to "**name a discovery that inspired you and then describe it in exactly six words**." Read a selection of their responses here. Follow NextGen Voices on Twitter with hashtag #NextGenSci. **—Jennifer Sills**

## NEUROSCIENCE

### Prechtl's general movements
Non-invasive neuro diagnosis.
Baby approved!
**Maria Lucrecia Cuneo,** Ciudad Autónoma de Buenos Aires, Argentina

### Columnar organization of the neocortex
One algorithm to rule them all
**Guthrie Phineas Dyce,** Canberra, ACT, Australia

### The neuron
Ramón y Cajal's
"enigmatic thought cell"
**Matheus de Castro Fonseca,** Pasadena, CA, USA

## IMMUNOLOGY

### Monoclonal antibodies
The immune system's ninja assassins, weaponized
**Xuan Zhang,** Guangdong, China

### Autoantibodies
The pirates of the immune system
**Ranjeet Singh Mahla,** Oxford, UK

### Immune basis of allograft rejection
Graft acceptance or rejection?
Host's choice!
**Jen Li,** Westmead, NSW, Australia

## INFECTIOUS DISEASE

### Ebola vaccine
Like cure in "Outbreak" movie, but real
**Gina Pontes,** Capital Federal, Buenos Aires, Argentina

### Smallpox vaccine
Jenner's vaccine: Cowpox's triumph, smallpox's defeat
**Yanpei Mai,** Guangzhou, Guangdong Province, China

### Rinderpest vaccine
One shot to eradicate a plague
**Anke Brüning-Richardson,** Yorkshire, UK

### Pandemic disease triangle
Disease: Intersection of microbe, environment, immunity
**Gabriela Rios-Sotelo,** Portland, OR, USA

## BIOLOGY

### Histopathology and immunohistochemistry
Rainbows under the microscope allow diagnosis!
**Melissa T. Duong,** Philadelphia, PA, USA

### Epimorphic regeneration
Salamanders regrow amputated limbs. Can humans?
**Leonardo M. R. Ferreira,** Charleston, SC, USA

### Adaptation to hypoxia
Normal hemoglobin level among Sherpa population
**Matiram Pun,** Calgary, AB, Canada

### Polymerase chain reaction
Technique to amplify DNA; thanks, bacteria!
**Anurup Mohanty,** Bangalore, Karnataka, India

## GENETICS

### Mendelian randomization
Nature's RCTs are cheaper than humans'!
**Haoxian Tang,** Guangdong, China

### CRISPR-Cas genome base editors
Editing DNA like a Word document
**Jasleen Gill,** Media, PA, USA

### Genetic influence on eye movements
Where I look = who I am
**Ben de Haas,** Giessen, Germany

## ECOLOGY

### Virtual water
Commodity trading: Equivalent to water diversion!
**Yuan Zhi,** Guiyang, Guizhou, China

### Belowground kin recognition
Plants:
Where family roots run deep!
**Harihar Jaishree Subrahmaniam,** Aarhus, Denmark

### Transgenic plants
Playing God?
I'm just improving crops!
**Mateo Daniel Ordoñez Rivera,** Quito, Pichincha, Ecuador

### Antifungal dual-use
Antifungals heal;
but also, fungicides kill.
**Norman van Rhijn,** Manchester, Lancashire, UK

## SOCIAL SCIENCE

### Social discounting
Generous? Maybe.
But every choice is biased.
**Fernanda S. Oda,** Lawrence, KS, USA

### Gender role congruity theory
Gender roles shape leaders' perceived competence.
**Lena Rudat,** Mannheim, Germany

### Problem-based learning
Code error fixed, lessons acquired: Jackpot!
**Ming Cheng,** Hefei, Anhui, China

## CHEMISTRY

### Ciamician's sunlight photochemistry
From the rooftops to the benchtops
**Terry McCallum,** Ottawa, ON, Canada

### Sabatier principle
Optimal catalysis:
Neither strong nor weak
**Hao Zhang,** Oxford, Oxfordshire, UK

### Fusion ignition
Limited energy to unlock limitless energy
**Rishi Jai Patel,** Philadelphia, PA, USA

### Field-driven colloidal assembly
You can treat nanoparticles like LEGOs!
**Ahmed Al Harraq,** Princeton, NJ, USA

## PHYSICS

### Quantum computing
A cat-qubit appeared in my dreams.
**Ginevra Fulco,** Milan, Italy

### Quantum cosmology
Supersymmetric wave function of the universe
**Paulo Moniz,** Covilhã, Portugal

## ENGINEERING

### Generalized Nash equilibrium
Cutting cake's no piece of cake.
**Yuanxing Xia,** Nanjing, Jiangsu Province, China

### Mass customization
Economies of scale, tailored for everyone
**Adrian Tymes,** Mountain View, CA, USA

### Bruun rule
Sandy coasts under threat of erosion
**Xiangming Cao,** Nanjing, Jiangsu, China

## COMPUTER SCIENCE

### Tortured phrases detector
Detecting the footprints of paraphrased plagiarism
**Houcemeddine Turki,** Sfax, Tunisia

### Moravec's paradox
AI: "It's hard to fold laundry!"
**Ju Wen,** Chengdu, Sichuan, China

## GEOGRAPHY

### Arctic and Tibetan Plateau teleconnection
Cryospheric changes affect the whole world!
**Xin Wan,** Chaoyang District, Beijing, China

### Hu Line
Geographical dividing line for understanding China
**Bing Xue,** Shenyang, China

10.1126/science.adi8737

Image of bacteria with helically shaped appendages known as flagella, which enable cell movement. Flagellin, the protein that forms the core of flagella, plays a key role in how host cells respond—or do not respond—to pathogens and commensals.

PRIZE ESSAY

GRAND PRIZE WINNER

**Sara Clasen**

Sara Clasen received her undergraduate degree from the University of Arizona in Tucson and a PhD from Johns Hopkins School of Medicine in Baltimore, USA. She is currently a postdoctoral fellow in Ruth Ley's Department of Microbiome Science at the Max Planck Institute for Biology, Tübingen, Germany. Her research examines immune responses to commensal microbiota. **www.science.org/doi/10.1126/science.adi6265**

NOSTER | Science MICROBIOME PRIZE

MICROBIOME

# The sound of silence

## How flagellins from commensals evade innate immunity

*By* **Sara Clasen**

In 1926, the Journal of Pathology and Bacteriology published the immunological notes of a little-known diphtheria researcher, A. T. Glenny. In table XLIII of that publication, he reported the notable effect of potassium aluminum on the guinea pig immune response to diphtheria toxin (*1*). Animals injected with the toxin in the presence of aluminum had a much more robust antigenic response compared with those exposed to the toxin alone. Additional "adjuvants," consisting of heat-killed bacterial extracts, were found by others to exert similar immunogenic effects. Later, this phenomenon was described by the preeminent immunologist Charles Janeway as, "the immunologist's dirty little secret" (*2*). What Glenny had stumbled on, and Janeway brought to the fore, was the critical role of the innate immune system in generating adaptive immune responses. To understand how these responses develop requires deeper knowledge of how innate immune signaling is regulated, which we studied in the context of the microbiome.

It is now broadly accepted that strong adaptive responses require activation of innate immunity, and that many adjuvants are ligands to innate immune receptors (*3*). The latter are frequently described as pattern recognition receptors because they evolved to bind highly conserved molecular patterns expressed by pathogens. But here is another little secret: These motifs are also produced by trillions of microbes that colonize the human gut, the vast majority of which are nonpathogenic and even beneficial to host health. How do innate immune receptors avoid reacting to these ligands?

The Ley laboratory focuses on Toll-like receptor 5 (TLR5) and its ligand, flagellin,

Department of Microbiome Science, Max Planck Institute for Biology, Tübingen, Germany.
Email: sara.clasen@tuebingen.mpg.de

CREDITS: (ILLUSTRATION) KATERYNA KON/SCIENCE SOURCE; (PHOTO) BRIGITTE SAILER/MAX PLANCK INSTITUTE FOR BIOLOGY

because their interaction exemplifies this conundrum. TLR5 is expressed by epithelial and immune cells, and both pathogenic and nonpathogenic bacteria produce flagellin proteins to build filaments for locomotion. When these filaments break down, TLR5 binds flagellin and initiates a pro-inflammatory response. Although the interaction between TLR5 and *Salmonella* FliC (the flagellin protein coded by *Salmonella*'s *fliC* gene) is well studied, relatively little is known about how the receptor responds to flagellins from nonpathogens, or "commensals." We analyzed both human gut metaproteome and metatranscriptomes from individuals without known disease to identify relevant, commensal-derived flagellins.

At both the protein and transcript level, we found that flagellins in the human gut are made almost exclusively by members of the Lachnospiraceae family (*4*). This taxonomic group is prevalent in human microbiomes across the globe with well-established roles in promoting host health (*5*). We wondered whether their flagellins induce innate immune responses similar to those caused by FliC.

To address this question, we characterized the interaction between TLR5 and 40 commensal-derived flagellins that are abundant in human gut metagenomes (*4*). We developed a TLR5-flagellin binding assay to infer recognition and measured nuclear factor κB (NF-κB) signaling in reporter cells to assess TLR5 activity. Recognition correlates with activation for most of the candidates. But for ~25%, receptor recognition is decoupled from activation—the flagellins bind the receptor but fail to induce the pro-inflammatory response.

We called this response "silent recognition" by TLR5 and explored how these "silent flagellins" are seen—but ignored—by innate immunity. The silent flagellin *Rh*FlaB, produced by the common gut resident *Roseburia hominis*, showed the strongest binding to TLR5 in our assay, and we used it to further interrogate how the TLR5-silent flagellin interaction differs from that of TLR5-FliC. Most silent flagellins are nearly identical to FliC in regions characterized as crucial for TLR5 activation (*6*). But when we mutated the primary interaction site in flagellin, we saw different effects on binding to TLR5: The *Rh*FlaB mutant showed a complete loss of binding, whereas the FliC mutant showed no change relative to the wild type (*4*). The resilience of the FliC mutant in binding TLR5 indicated the presence of a secondary, allosteric binding site in FliC that was missing from *Rh*FlaB. We mapped its location then explored its effect by transplanting it into silent flagellins. As expected, the allosteric binding site restored the *Rh*FlaB mutant's ability to bind TLR5. And unexpectedly, it made wild-type *Rh*FlaB orders of magnitude better at activating TLR5.

How does this binding site exert such a large impact on TLR5 signaling? For answers, we looked to the mechanism behind TLR5 activation. Flagellin is thought to induce receptor dimerization because TLR5 is a homodimer in its active state (*7*). Therefore, we explored whether the allosteric binding site facilitates this process. The effort bore unexpected fruit: Although we did not detect FliC-induced dimerization, we did discover that a substantial fraction of TLR5 on the cell surface exists as a dimer in the absence of flagellin (*4*). The FliC allosteric site mediates binding to these preformed dimers. Silent flagellins, consequently, cannot bind dimeric TLR5 unless they are equipped with the allosteric site from FliC. Thus, evasion of TLR5 dimers enables silent flagellins to avoid activating the receptor.

Our discovery of silent flagellins illustrates one way by which innate immune receptors tolerate ligands from commensals. To contextualize these findings, we explored the prevalence of silent flagellins in the human gut. We screened additional flagellins from nearly a hundred commensal bacterial species to expand our list of silent flagellins, then measured their abundance in more than 1700 human gut metagenomes from around the world. Our analysis revealed that silent flagellins are ubiquitous. However, their levels are substantially lower in populations from industrialized regions (*4*). The factors behind this loss are unclear, as are the consequences for human health.

What is increasingly apparent is that Lachnospiraceae flagellins can become the targets of antigenic responses, and these responses are overrepresented in disease states, including inflammatory bowel disease (IBD) (*8*, *9*). The antibodies elevated in IBD target flagellins that strongly stimulate TLR5, as well as silent flagellins. However, TLR5's role in this process is not well understood. Notably, the prevalence of IBD is much higher in industrialized countries (*10*) and correlates with reduced silent flagellin abundance in these populations.

Since Janeway's remarks in 1989, many insights have been made into how innate immunity enhances adaptive responses. Although vaccine adjuvants have long shown how this relationship can be harnessed for public health, antigenic responses to commensal microbiota indicate that it can also be detrimental. Our continued work on the interaction between innate immune receptors and the microbiome—using methods from both computational and molecular biology (*11*)— is critical to understand how these maladaptive responses develop and to mitigate their effects. ■


**REFERENCES AND NOTES**

1. A. T. Glenny, C. G. Pope, H. Waddington, U. Wallace, *J. Pathol. Bacteriol.* **29**, 31 (1926).
2. C. A. Janeway Jr., *Cold Spring Harb. Symp. Quant. Biol.* **54**, 1 (1989).
3. R. L. Coffman, A. Sher, R. A. Seder, *Immunity* **33**, 492 (2010).
4. S. J. Clasen *et al.*, *Sci. Immunol.* **8**, eabq7001 (2023).
5. P. Louis, H. J. Flint, *FEMS Microbiol. Lett.* **294**, 1 (2009).
6. K. D. Smith *et al.*, *Nat. Immunol.* **4**, 1247 (2003).
7. S.-I. Yoon *et al.*, *Science* **335**, 859 (2012).
8. K. L. Alexander *et al.*, *Gastroenterology* **161**, 522 (2021).
9. T. Vogl *et al.*, *Sci. Adv.* **8**, eabq2422 (2022).
10. G. G. Kaplan, J. W. Windsor, *Nat. Rev. Gastroenterol. Hepatol.* **18**, 56 (2021).
11. S. J. Clasen, W. Shao, H. Gu, P. J. Espenshade, *eLife* **6**, e28563 (2017).


10.1126/science.adi6265

**FINALIST**

## Christopher J. Stewart

Christopher Stewart received undergraduate degrees and a PhD from Northumbria University. After completing his postdoctoral fellowship at Baylor College of Medicine, he started his laboratory in the Translational and Clinical Research Institute at Newcastle University in 2018. His research is focused on microbial-host interaction in the gut of infants born prematurely, using multi-omic analysis of clinical samples in combination with basic microbiology and an experimental organoid coculture system, to develop disease biomarkers and targeted therapeutic interventions to promote health. **www.science.org/doi/10.1126/science.adi6318**

**FINALIST**

## Christoph A. Thaiss

Christoph A. Thaiss received his undergraduate degrees from the University of Bonn, Yale University, and ETH Zürich and a PhD from the Weizmann Institute of Science. After completing graduate school, he started his laboratory in the Microbiology Department of the Perelman School of Medicine at the University of Pennsylvania in 2018. His research studies the multifaceted interactions between environmental factors, the gut microbiome, the immune system, metabolism, and the brain. **www.science.org/doi/10.1126/science.adi6329**

PRIZE ESSAY

FINALIST
**Christopher J. Stewart**

Christopher Stewart received undergraduate degrees and a PhD from Northumbria University. After completing his postdoctoral fellowship at Baylor College of Medicine, he started his laboratory in the Translational and Clinical Research Institute at Newcastle University in 2018. His research is focused on microbial-host interaction in the gut of infants born prematurely, using multi-omic analysis of clinical samples in combination with basic microbiology and an experimental organoid coculture system, to develop disease biomarkers and targeted therapeutic interventions to promote health. **www.science.org/doi/10.1126/science.adi6318**

NOSTER | Science MICROBIOME PRIZE

PHOTO: STEVEN O'GORMAN

MICROBIOME

# Diet-microbe-host interaction in early life

## Breast milk bioactives are important for infant microbiome and immunity

*By* **Christopher J. Stewart**

Breast milk is a complex biofluid crafted over millions of years of evolution to provide optimal nutrition along with other bioactive compounds to the infant. Human milk oligosaccharides (HMOs) are complex sugars that represent the fourth most abundant component of breast milk, which makes them more numerous than proteins and other vital nutrients, yet they provide no direct nutritional benefit to the infant (*1*). Human evolution does not favor wasteful processes, which suggests that HMOs play another crucial role in infant development. Recent work has borne this out and highlighted HMOs' ability to act as prebiotics. To improve human health throughout life by means of targeted intervention, it is paramount to better understand the three-way interaction between diet, microbe, and host.

The environmental determinants of diabetes in the young (TEDDY) study collected over 12,000 longitudinal stool samples from more than 900 infants and found that receipt of breast milk was the most important factor shaping the infant gut microbiome throughout the first year of life, after which breastfeeding seldom occurred (*2*). Taxonomic analysis showed that *Bifidobacterium* spp. were closely linked to receipt of breast milk, which followed on previous work that described the distinctive genetic capacity of *Bifidobacterium* species to utilize HMOs from breast milk (*3*).

Such work has been performed in the context of infants born at full term, but a major research gap relates to diet-microbe-host interaction in infants born prematurely. This important population has been the focus of my group's recent work, which revealed that HMOs modulate the preterm infant gut microbiome, increase *Bifidobacterium* colonization, and reduce disease risk. Further, probiotics provide the means to deliberately modulate the preterm gut microbiome in a highly specific and reproducible way.

Unlike infants born at full term, extremely preterm infants (i.e., <32 weeks' gestation) have an immature intestinal architecture and an underdeveloped immune system. They are also less likely to be breastfed and receive limited exposure to microbes during the first months of life, which leads to a reduction in potentially beneficial bacteria. This reduction has been implicated in the pathogenesis of necrotizing enterocolitis (NEC), a serious inflammatory-mediated bowel disease that represents the leading cause of death in this population (*4*). Although the risk factors for NEC are complex, the most protective factor against disease for these infants is unequivocally receiving their mother's own breast milk.

Because the preterm gut can become leaky, translocation of microbes into the bloodstream and/or intestinal cell death represent major problems. However, evidence from my research and others has shown that certain types of bacteria, most notably HMO-utilizing *Bifidobacterium*, may increase gut and immune maturation (*5*–*8*). Other work has also suggested a possible role of specific HMOs in protection against NEC (*9, 10*).

We recently expanded upon this research in our investigation of HMOs in mothers' milk and longitudinal infant gut microbiome samples in 33 NEC cases and 37 matched controls. HMO profiling revealed that a single HMO, disialyllacto-*N*-tetraose (DSLNT), was lower in breast milk received by infants who went on to develop NEC and could predict disease with high accuracy (*11*). We used univariate modeling to determine an optimal DSLNT threshold for disease risk and modeled how the developing preterm gut microbiome was affected by receipt of DSLNT above or below that threshold. Infants that received milk with DSLNT below the threshold were associated with slower microbiome maturation and reduced colonization by preterm gut community types (PGCTs) dominated by *Bifidobacterium* species. Specifically, healthy babies who received adequate DSLNT showed quicker progression to potentially

Translational and Clinical Research Institute, Newcastle University, Newcastle, UK.
Email: christopher.stewart@newcastle.ac.uk

beneficial PGCT-4 (dominated by *B. longum* and *B. bifidum*) and PGCT-5 (dominated by *B. breve*).

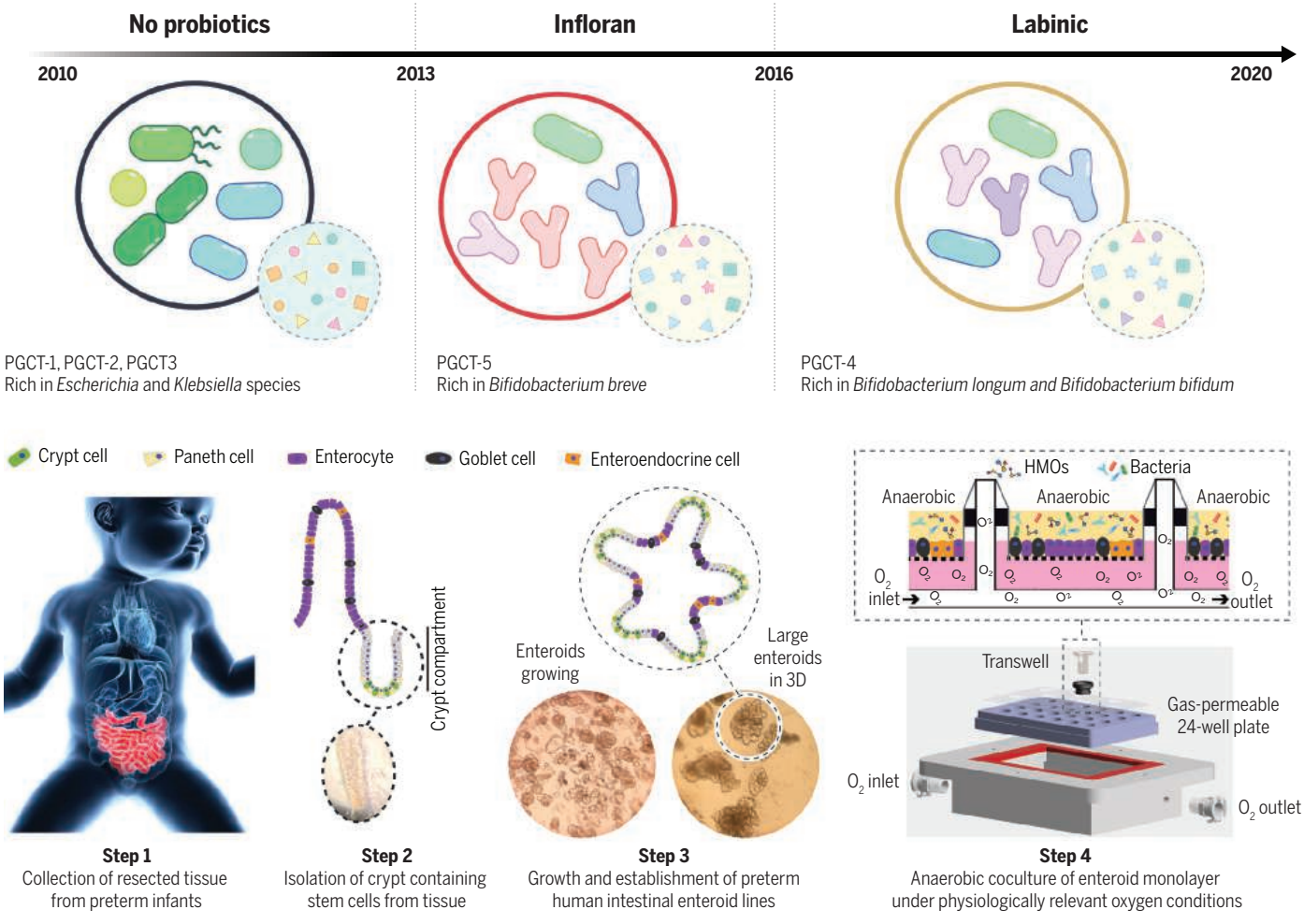
Given these important links between HMOs in breast milk and PGCT colonization, we asked if different probiotic products could be used to effectively tailor preterm gut microbiome development (*12*). We leveraged the Great North Neonatal Biobank, which has collected daily samples since 2010 to enable research into the preterm gut microbiome and immune development. In this time, the neonatal intensive care unit (NICU) at Great North Children's Hospital in Newcastle, UK, started using probiotics, initially Infloran in 2013 and then Labinic in 2016. Although the two probiotic products included the same species (*B. bifidum* and *Lactobacillus acidophilus*), with the addition of *B. longum* in Labinic, the strains for each product were distinct. This difference in strains led to consistent and important impacts on the preterm gut microbiome, such that infants that received Labinic quickly transitioned to PGCT-4 and infants receiving Infloran transitioned to PGCT-5. Infants in the NICU before probiotics were introduced lacked colonization by *Bifidobacterium* and instead were dominated by pathobionts including *Escherichia* and *Klebsiella* (see the figure).

Although associations derived from "omic" analysis of clinical samples provide important insights, it is not possible to determine causality or underlying mechanisms. We recently described a protocol for the generation of human intestinal organoids from preterm gut tissue (*13*) and engineered an intestinal organoid physoxic co-culture (IOPC) system that allows simultaneous growth of anaerobic bacteria and intestinal organoids (*14*) (see the figure). We then used the IOPC system to demonstrate that preterm ileum-derived organoids have distinct gene expression when compared to adult ileum-derived organoids (*15*).

Using the preterm intestinal-derived organoid model revealed critical links between PGCTs, namely, that PGCT-4 and PGCT-5 had distinct host epithelial transcription when compared to the other PGCTs and media-only control (*12*). Metabolomic analysis of stool samples also found that PGCT-4 and PGCT-5 were more comparable functionally, despite differences in microbial taxonomy, and were distinct from the nonprobiotic PGCTs 1–3 (see the figure). Such research has important implications for the design of preterm infant probiotics and ultimately leads one step closer to per-

## Overview of preterm work spanning descriptive microbiome sequencing, functional metabolomic profiling, and mechanistic organoid coculture

(Top) Timeline of probiotic use in preterm infants reveals a strain-specific impact on taxonomic composition. Large circles represent bacterial taxonomy, and small circles represent the functional metabolites in preterm infant stool. (Bottom) Schematic representation of the process of establishing, expanding, and experimenting with preterm intestinal-derived organoids.




PGCT, preterm gut community type; HMOs, human milk oligosaccharides.

sonalized medicine for gut health.

Further work is ongoing to validate DSLNT as a biomarker for disease and to understand the potential mechanism to enable targeted therapeutic intervention using HMOs and HMO-utilizing probiotics. Better understanding of the diet-microbe-host interaction holds incredibly exciting possibilities to predict, diagnose, and manipulate the microbiome of preterm infants to reduce disease risk.


#### REFERENCES AND NOTES

1. A. C. Masi, C. J. Stewart, *iScience* **25**, 103542 (2022).
2. C. J. Stewart *et al.*, *Nature* **562**, 583 (2018).
3. D. A. Sela *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18964 (2008).
4. A. C. Masi, C. J. Stewart, *Early Hum. Dev.* **138,** 104854 (2019),
5. C. J. Stewart *et al.*, *Microbiome*. **4**, 67 (2016).
6. C. J. Stewart *et al.*, *Microbiome*. **5**, 75 (2017).
7. S. Fanning *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2108 (2012).
8. C. J. Stewart, *Nat. Microbiol*. **6**, 1335 (2021).
9. E. Jantscher-Krenn *et al.*, *Gut* **61**, 1417 (2012).
10. C. A. Autran *et al.*, *Gut* **67**, 1064 (2018).
11. A. C. Masi *et al.*, *Gut* **70**, 2273 (2021).
12. L. C. Beck *et al.*, *Nat. Microbiol*. **7**, 1525 (2022).
13. C. J. Stewart, M. K. Estes, S. Ramani, *Methods Mol. Biol.* **2121**, 185 (2020).
14. T. Y. Fofanova *et al.*, bioRxiv 555755 (2019).
15. A. C. Masi *et al.*, *Gut* 10.1136/GUTJNL-2021-326552 (2021).



#### ACKNOWLEDGMENTS

C.J.S. is supported by the Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant 221745/Z/20/Z), a Newcastle University Academic career Track (NUAcT) Fellowship, and the 2021 Lister Institute Prize Fellow Award. He has performed consultancy for Astarte Medical and received lecture honoraria from Nestle Nutrition Institute. L.Beck (Newcastle University) assisted with figure design.

10.1126/science.adi6318

PRIZE ESSAY

FINALIST

**Christoph A. Thaiss**

Christoph A. Thaiss received his undergraduate degrees from the University of Bonn, Yale University, and ETH Zürich and a PhD from the Weizmann Institute of Science. After completing graduate school, he started his laboratory in the Microbiology Department of the Perelman School of Medicine at the University of Pennsylvania in 2018. His research studies the multifaceted interactions between environmental factors, the gut microbiome, the immune system, metabolism, and the brain.

www.science.org/doi/10.1126/science.adi6329

MICROBIOME

# A microbiome exercise

## Gut-brain connections drive the motivation to work out

*By* **Christoph A. Thaiss**

Exercise is arguably the most potent and attainable lifestyle factor that individuals can use to safeguard themselves against a large variety of diseases, including metabolic, cardiovascular, neurodegenerative, and neoplastic diseases (*1*). Despite its vast health benefits, only a small fraction of the population adheres to the World Health Organization's recommendation of 150 to 300 minutes of moderate-intensity exercise per week (*2*). Busy schedules and rainy weather aside, what are the physiological bottlenecks that prevent us from engaging in regular physical activity?

My research team at the University of Pennsylvania, spearheaded by PhD student Lenka Dohnalová, set out to answer this question using a large cohort of genetically diverse outbred mice. We observed marked interindividual variability in exercise capacity, with the top performers outrunning the weakest animals by more than 10-fold (*3*). We then carried out deep molecular and physiological profiling of these mice and used a machine-learning approach to nominate individual measurements that predict treadmill and wheel running. To our surprise, elements of the gut microbiome topped the list. We thus probed the functional importance of the microbiome in exercise performance and carried out ergometric measurements under gnotobiotic conditions (i.e., with defined microbial colonization). We noted that the absence of the microbiome in germ-free or antibiotics-treated mice diminishes voluntary and endurance exercise capacity by ~50%. We determined that this effect of the microbiota was driven by certain taxonomic elements, including members of the Lachnospiraceae family. These experiments suggest that the gut microbiome may be an important contributor to the interindividual variability in physical activity.

The question of how the microbiome affects exercise performance led us to an intriguing discovery. Contrary to our expectations, the effects of intestinal microbes on physical performance were not mediated by tissues classically associated with exercise, such as skeletal muscle. Rather, the microbiome appeared to fundamentally influence the impact of exercise on the brain. Intense physical activity stimulates numerous neurochemical responses in the brain, including a surge in dopamine release in the striatum (*4*). And interestingly, in the absence of the microbiome, this surge was severely diminished.

The release of striatal dopamine is known to cause feelings of pleasure, motivation, and reward (*5*)—all elements of the "runner's high" phenomenon that some individuals experience after intense exercise. On the basis of these findings, we speculated that the microbiome may enhance exercise capacity through its ability to modulate exercise-induced dopamine responses. Through a series of pharmacological and chemogenetic interventions, we found that dopamine deficiency was indeed the cause for reduced exercise in microbiome-depleted animals. Consistently, restoring dopamine signaling fully restored their running ability (*3*).

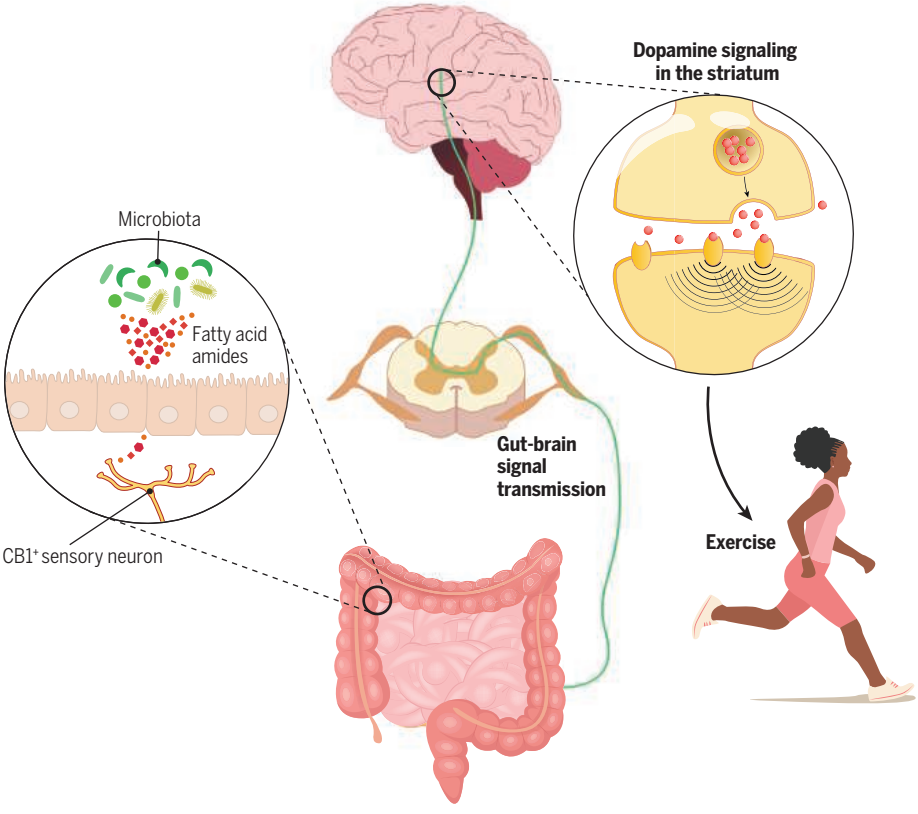
How do microbes that reside in the gastrointestinal tract control neurotransmitter levels in the brain? The communication between the microbiome and the brain can occur either through microbial molecules entering the systemic circulation or through direct neuronal connections between the gut and brain (*6*). We did not find evidence for a microbiome-derived systemic mediator that influences exercise capacity. By contrast, inhibiting gut-innervating sensory neurons recapitulated the effect of microbiome depletion on exercise, whereas stimulation of sensory afferents restored exercise capacity even in the absence of the microbiome.

To determine the mechanisms by which the microbiome influences the activity of sensory neurons during exercise, we combined microbial genetics, metabolomics, and neuronal recordings. We identified a pathway whereby microbial production of fatty acid amides (FAAs) enhanced sensory neuron activity through the endocannabinoid receptor CB1. This increase in neuronal activity, in turn, boosted striatal dopamine signaling during exercise (see the figure). Intriguingly, this pathway can be harnessed to modulate exercise performance through gastrointestinal interventions. For instance, colonizing germ-free mice with genetically engineered FAA-producing bacteria or feed-

Microbiology Department, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
Email: thaiss@pennmedicine.upenn.edu

## A meta-organismal pathway regulates the motivation for exercise

Microbiota-derived fatty acid amides are sensed by the endo-cannabinoid receptor CB1 on gut-innervating neurons, which enhances their activity during exercise. Sensory neuron activation, in turn, augments dopamine signaling in the striatum and thereby drives exercise performance.




ing antibiotics-treated mice an FAA-containing diet restored their exercise capacity. However, when we used a peripheral CB1 inhibitor or a central dopamine receptor blocker, the positive effects of these interventions were negated (*3*).

Our finding has several important implications. It suggests that gastrointestinal signals can serve as a regulating factor on the activity of exercise-responsive neurons in the brain, potentially enhancing the generation of the "runner's high" feeling. The evolutionary purpose of this gut-brain pathway is yet to be determined, but we speculate that it may link an animal's decision to engage in prolonged physical activity to the nutritional status of the gastrointestinal tract.

The mental state plays a crucial role in exercise participation and performance, both in elite athletes and hobby runners (*7*). Our discovery of a microbiome-dependent pathway may contribute to a better understanding of the factors that motivate some individuals to reach their peak performance, whereas others prefer to observe from the sidelines. The malleability of the microbiome and its metabolites offers a potential avenue to increase exercise engagement and its health benefits globally.

Striatal dopamine is a critical mediator of motivated behavior beyond just exercise. Our discovery opens up the possibility of leveraging this pathway to influence other dopamine-dependent brain functions, such as learning, mood, and addiction. Further explorations into the vast universe of microbiome-derived molecules that influence neuronal activity will provide valuable insights into the potential of shaping the mind through the gastrointestinal tract.

### REFERENCES AND NOTES


1. P. D. Neufer *et al.*, *Cell Metab.* **22**, 4 (2015).
2. D. L. Blackwell, T. C. Clarke, *Natl. Health Stat. Report*, 1 (2018).
3. L. Dohnalova *et al.*, *Nature* **612**, 739 (2022).
4. D. M. Friend *et al.*, *Cell Metab.* **25**, 312 (2017).
5. J. D. Berke, *Nat. Neurosci.* **21**, 787 (2018).
6. G. Agirman, K. B. Yu, E. Y. Hsiao, *Science* **374**, 1087 (2021).
7. P. J. Teixeira, E. V. Carraca, D. Markland, M. N. Silva, R. M. Ryan, *Int. J. Behav. Nutr. Phys. Act.* **9**, 78 (2012).


### ACKNOWLEDGMENTS


I am indebted to the members of my research team and our collaborators who make discoveries like the one described in this essay possible. They embody the excitement, camaraderie, and joy of the scientific endeavor.

10.1126/science.adi6329


GRAPHIC: THAISS *ET. AL*

# RESEARCH

## IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

BIOGEOGRAPHY

### When continents collide

The movement of species across newly connected continents millions of years ago still shapes flora and fauna today. Skeels *et al.* showed that species' dispersal ability, climate tolerances, and the climate in which they evolved help to explain why biotic exchanges are typically unequal, with more species spreading from one continent than the other. Using a model simulating species ranges and diversification paired with paleoenvironmental reconstructions, the authors found that precipitation tolerance influenced vertebrate species movements across Wallace's Line, which separates the distinct biota of Australia and New Guinea from that of Southeast Asia. Species that evolved in dry Australia were less able to cross to Asia, whereas the swath of tropical forest across the region allowed more species to move in the other direction through New Guinea. —BEL

*Science*, adf7122, this issue p. 86

The Sulawesi bear cuscus (*Ailurops ursinus*) is, in fact, a marsupial species endemic to the region around Wallace's Line.

PHOTO: PHILIPPE CLÉMENT/ARTERRA/UNIVERSAL IMAGES GROUP VIA GETTY IMAGES

BATTERIES

### Rapid ionic transport in inorganic solids

Solid electrolytes have been developed to allow solid-state batteries to compete with liquid electrolyte batteries, but there is still a challenge in ensuring good contact with the electrodes, which thus limits the thickness of electrodes that can be used in practice. Li *et al*. applied the principles of high-entropy materials to improve upon the properties of existing electrolytes through the sensible partial substitution of addition elements. The authors demonstrate the incorporation of their materials into a cathode structure, which provides the high ionic conductivity required for solid-state battery applications. —MSL

*Science*, add7138, this issue p. 50

STRUCTURAL BIOLOGY

### Snapshots of ribosome dynamics in human cells

Ribosomes are molecular machines that translate genetic information into proteins. Their active site is the target of many small molecules, including several antibiotics and the anticancer drug homoharringtonine. Although structural biologists have studied how small molecules interfere with ribosome function in the test tube, Xing *et al*. now reveal structures of translating ribosomes within human cells using cryo–electron tomography. The authors visualized translation intermediates at high resolution and elucidated in detail how homoharringtonine acts on ribosomes. This work marks a technical advance in solving high-resolution structures from inside human cells and adds to our understanding of this drug's effect on cells. —SMH

*Science*, adh1411, this issue p. 40

ORGANIC CHEMISTRY

### Put a ring in it

Traditionally, chemists make complex molecules by first assembling the central framework and then modifying the periphery. More recently, however, there has been an increased focus on late-stage framework editing by inserting or removing individual atoms. Wang *et al*. now report the insertion of a whole ring into a preexisting ring. Specifically, they relied on a photoredox catalyst to oxidize thiophene derivatives, activating them for the insertion of highly strained bicylobutane.

The bicyclic products feature an unusual eight-membered ring motif. —JSY

*Science*, adh9737, this issue p. 75

MOLECULAR BIOLOGY

## Kissing and kicking in transcription

The spatial organization of chromatin and specific factors within the nucleus regulates gene expression. To visualize nuclear organization at the molecular scale in developing embryos, Pownall *et al*. developed chromatin expansion microscopy (ChromExM), which drastically improves the resolution of light microscopy by physically expanding embryos embedded in a series of swellable hydrogels without distorting chromatin organization (see the Perspective by Stasevich and Kimura). This procedure enabled the visualization of individual nucleosomes and revealed how the pioneer factor Nanog interacts with them. The authors also observed RNA polymerase II nanostructures engaged in active transcription, and developed a "kiss and kick" model of enhancer-promoter interactions in which transcription elongation leads to the physical separation of enhancers and promoters. —DJ

*Science*, ade5308, this issue p. 92; see also adi8187, p. 26

QUANTUM SIMULATION

## Watching fermions pair

Ultracold atoms in optical lattices can be used to gain insight into the behavior of interacting quantum matter. Some of the most general physics that can be simulated is accessed through the Fermi-Hubbard model, which is relevant to many solid-state systems. Hartke *et al*. placed attractively interacting fermionic atoms of potassium into an optical lattice and used a quantum microscope to image the system and measure correlations between the lattice sites. Their results suggest the formation of nonlocal atomic pairs with a size that depends on the strength of attraction. —JS

*Science*, ade4245, this issue p. 82

CELL CYCLE

## RhoGEF hitches a ride in dividing cells

In cytokinesis, the final phase of cell division, the plasma membrane progressively pinches inward along the cell equator. Warecki and Tao investigated how the activation of the enzyme RhoA is precisely localized to enable this process. In a purified protein system, the activator of RhoA, RhoGEF, accumulated at microtubule ends due to a motor protein complex called centralspindlin. In cells, this accumulation of RhoGEF sustained RhoA activity and cytokinesis progression at the cell equator. The findings support a long-held hypothesis for how cells locally activate RhoA during cytokinesis. —LKF

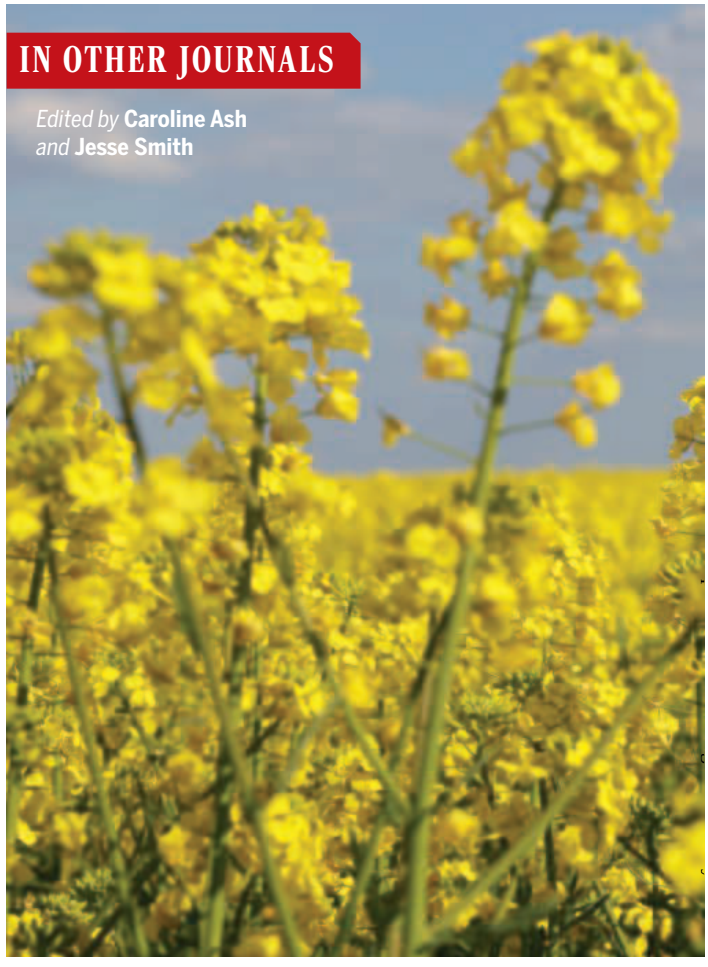
*Sci. Signal.* (2023) 10.1126/scisig.eadh0601

HIV

## bNAbs for babies

Broadly neutralizing antibodies (bNAbs) have shown promise as a companion or alternative to antiretroviral therapy (ART) for HIV-1, although their efficacy in the pediatric population remains unclear. Shapiro *et al.* report the results of a prospective clinical trial in which children on ART from birth in Botswana were treated with two HIV-1 bNAbs. The treatment was administered initially in combination with ART and was then given alone. Eleven of 25 infants maintained viral suppression during the bNAb-only step, and the authors found that these infants had more favorable HIV-1 reservoir characteristics, including a smaller initial proviral reservoir and susceptibility of those proviruses to bNAb neutralization. These results highlight the potential of bNAb treatment for infants and children living with HIV-1. —CSM

*Sci. Transl. Med.* (2023) 10.1126/scitranslmed.adh0004




CONSERVATION

## The human predator

Predators are formative components of ecosystems. Although modern humans tend not to think of ourselves as predators, ecologically, we are the predator with the broadest and largest impact across species. Darimont *et al.* surveyed more than 45,000 vertebrate species and estimated that over a third are "preyed" upon by humans through hunting and the pet trade, among other activities. Removal of these species by humans may have ecosystem-level impacts, as 40% of harvested species are threatened by our exploitative consumption. —SNV

*Commun. Biol.* (2023) 10.1038/s42003-023-04940-w

PREDATION

## Lesser fleas ad infinitum

We understand strategic aspects of the binary lifestyles of bacteria chiefly from model species. However, bacteria fall prey to smaller bacteria, the growth and division of which may be regulated in different ways. Using single-cell methods, Santin *et al.* systematically explored the life cycle of the bacterial predator *Bdellovibrio bacteriovorus*. This predator forms nonbinary replicative filaments within compartments in bacterial prey cells, which generate variable numbers of daughter cells. Larger prey cells mean more offspring than smaller cells, but it is the nutritional quality of the prey (experimentally manipulated through the growth medium) that ultimately determines the growth rate of the predator. That the timing of key cellular processes scales with the size and quality of prey cells indicates that this tiny predator can adjust its life cycle to the variability of the prey it encounters. —CA

*Curr. Biol.* (2023) 10.1016/j.cub.2023.04.059

## ALSO IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

MICROBIAL ECOLOGY

### Leaf microbiome interactions

Plant surfaces host complex microbial communities. Carbon source composition dictates whether microbes can live there, but species interactions make it difficult to predict which microbe combinations will thrive on a given substrate. Schäfer *et al*. examined the carbon source preferences of 224 bacterial strains from plant leaves and used metabolic genome-scale models to simulate pairwise interactions. Most interactions were found to be negative, with bacteria growing less well in pairs than in isolation. Metabolic flux analyses indicated that occasional positive interactions may occur for versatile strains that use amino acids and organic acids more easily. This work may inform the construction of synthetic microbial communities according to substrate. —MRS

*Science*, adf5121, this issue p. 42

STRUCTURAL BIOLOGY

### Guide RNA editing

In the parasitic protozoan *Trypanosoma brucei*, most mitochondrial genes are corrupted, and their transcripts must be repaired to restore coding capacity. Directed by guide RNAs (gRNAs), a structure called the editosome executes a cascade of messenger RNA (mRNA) cleavage, uridine insertions and deletions, and religation. Using cryo–electron microscopy and molecular approaches, Liu *et al*. visualized states of the substrate-binding complex responsible for gRNA stabilization and gRNA-mRNA duplex formation. This work provides a detailed framework for understanding mRNA editing by the editosome. —DJ

*Science*, adg4725, this issue p. 43

STRUCTURAL BIOLOGY

### Making the human ribosome large subunit

Human ribosome biogenesis is a highly regulated process involving a vast ensemble of assembly factors that orchestrate the formation of both ribosomal subunits. The mechanisms underlying the early assembly of the human large ribosomal subunit (pre-60*S*) remain elusive. To provide insights into human ribosome assembly, Vanden Broeck and Klinge used a combination of genome editing, cryo–electron microscopy, and functional studies to reveal the structures of human pre-60*S* particles as they matured in the nucleolus and nucleus. This collection of structures highlights how assembly factors ensure the unidirectional maturation of human large ribosomal subunit precursors. —DJ

*Science*, adh3892, this issue p. 44

FOREST PHENOLOGY

### Opposite warming effects by season

Global warming is changing the timing of photosynthesis, with leaves emerging earlier in spring in the temperate and boreal zones. A longer growing season could mean greater carbon sequestration in forests, but the timing of leaves falling in autumn depends on multiple cues, making it difficult to predict. Zohner *et al*. investigated how leaf senescence relates to day length, temperature, and early-season photosynthesis across northern forests using remote sensing, ground observations, and experimental data. They found that warming had opposing effects on senescence dates depending on when it occurred: Warmer springs with higher photosynthesis correlated with earlier senescence, wheras warmer temperatures in autumn delayed senescence. Incorporating this shift in knowledge may improve predictions of vegetation response to climate change. —BEL

*Science*, adf5098, this issue p. 45

PHYSICS

### Sizing up the electron's dipole moment

The puzzling imbalance between matter and antimatter in the universe can be explained by the breaking of charge parity symmetry. The standard model of particle physics does predict a slight breaking of this symmetry but is insufficient to explain the observations. Many extensions to the standard model have been proposed to resolve this discrepancy. To test such model extensions, tabletop experiments that measure the electron's electric dipole moment (eEDM), a measure of symmetry breaking, have looked very promising. Roussy *et al*. exploited the large electric fields inside the polar molecular ions of hafnium fluoride to measure the eEDM to extremely high precision (see the Perspective by Fan and Jayich). The uncertainty of the measurement compares favorably to those achievable through accelerator-based experiments. —JS

*Science*, adg4084 this issue p. 46; see also adi8499, p. 28

PLANT SCIENCE

### Guard cell specification

Many tissue-patterning processes involve asymmetric cell divisions that determine subsequent cell fate. Working with the *Arabidopsis* guard cell lineage, Muroyama *et al*. found that the polarly localized proteins BASL/BRXf generate a region where cell division planes cannot occur (see the Perspective by Goldy and Caillaud). BASL/BRXf locally suppressed formation of the preprophase band, which then formed elsewhere to appropriately place the cell division site. When the preprophase band was disrupted, neither guard cells nor epidermal pavement cells formed. Within the polarity domain, organizational dynamics of microtubule plus-ends were disrupted, depleting microtubules from the zone. This study uncovers a mechanism for plant tissue patterning that ensures the inheritance of cell fate determinants to a single daughter cell. —MRS

*Science*, add6162, this issue p. 54; see also adi6664, p. 27

SOLAR CELLS

### Improving perovskite/silicon solar cells

Two studies show how interfaces between perovskite layers and silicon cells in tandem solar cells can be modified to improve performance (see the Perspective by De Wolf and Aydin). Mariotti *et al*. showed that an ionic liquid, piperazinium iodide, improved band alignment and enhanced charge extraction at the interface of a trihalide perovskite and a $C^{60}$ electron-transporting layer by creating a positive dipole. With these modifications, a 2.0-volt open circuit voltage was achieved in a silicon tandem cell. Chin *et al*. report the uniform deposition of the perovskite top cell on the micropyramids of crystalline silicon cells to achieve high photocurrents in tandem solar cells. Two different phosphonic acids improved the perovskite crystallization process and also minimized recombination losses. These modifications yielded perovskite/silicon tandem cells with certified power conversion efficiencies of more than 31% for active areas of at least 1 square centimeter. —PDS

*Science*, adg0091, adf5872 this issue p. 59, p. 63; see also adi6278, p. 30

IMMUNE REGULATION

### Pinning down KIR3DL3

Human killer immunoglobulin-like receptor (KIR) KIR3DL3 has been characterized as a

conserved but polymorphic inhibitory KIR that can recognize the protein HHLA2, which is up-regulated in tumors, as a ligand. Palmer *et al.* used high-resolution flow cytometry and single-cell transcriptomics to characterize KIR3DL3-expressing cells. They detected enriched expression of KIR3DL3 in γδ and $CD8^+$ T cells but not in natural killer cells. KIR3DL3-expressing cells were most frequently found in the lungs and digestive tract and were rare in peripheral blood and thymus. Blood $KIR3DL3^+$ T cells had an activated transitional memory phenotype and were considered hypofunctional. KIR3DL3 ligation could inhibit T cell receptor–mediated stimulation, and KIR3DL3 polymorphisms were most likely to reduce expression. KIR3DL3 expression appears to be up-regulated in parallel or downstream from T cell receptor stimulation, and individuals can have varied KIR3DL3 expression that can affect immune responses. —CNF

*Sci. Immunol.* (2023)
10.1126/sciimmunol.ade5343

The bicyclic products feature an unusual eight-membered ring motif. —JSY

*Science*, adh9737, this issue p. 75

MOLECULAR BIOLOGY

## Kissing and kicking in transcription

The spatial organization of chromatin and specific factors within the nucleus regulates gene expression. To visualize nuclear organization at the molecular scale in developing embryos, Pownall *et al*. developed chromatin expansion microscopy (ChromExM), which drastically improves the resolution of light microscopy by physically expanding embryos embedded in a series of swellable hydrogels without distorting chromatin organization (see the Perspective by Stasevich and Kimura). This procedure enabled the visualization of individual nucleosomes and revealed how the pioneer factor Nanog interacts with them. The authors also observed RNA polymerase II nanostructures engaged in active transcription, and developed a "kiss and kick" model of enhancer-promoter interactions in which transcription elongation leads to the physical separation of enhancers and promoters. —DJ

*Science*, ade5308, this issue p. 92; see also adi8187, p. 26

QUANTUM SIMULATION

## Watching fermions pair

Ultracold atoms in optical lattices can be used to gain insight into the behavior of interacting quantum matter. Some of the most general physics that can be simulated is accessed through the Fermi-Hubbard model, which is relevant to many solid-state systems. Hartke *et al*. placed attractively interacting fermionic atoms of potassium into an optical lattice and used a quantum microscope to image the system and measure correlations between the lattice sites. Their results suggest the formation of nonlocal atomic pairs with a size that depends on the strength of attraction. —JS

*Science*, ade4245, this issue p. 82

CELL CYCLE

## RhoGEF hitches a ride in dividing cells

In cytokinesis, the final phase of cell division, the plasma membrane progressively pinches inward along the cell equator. Warecki and Tao investigated how the activation of the enzyme RhoA is precisely localized to enable this process. In a purified protein system, the activator of RhoA, RhoGEF, accumulated at microtubule ends due to a motor protein complex called centralspindlin. In cells, this accumulation of RhoGEF sustained RhoA activity and cytokinesis progression at the cell equator. The findings support a long-held hypothesis for how cells locally activate RhoA during cytokinesis. —LKF

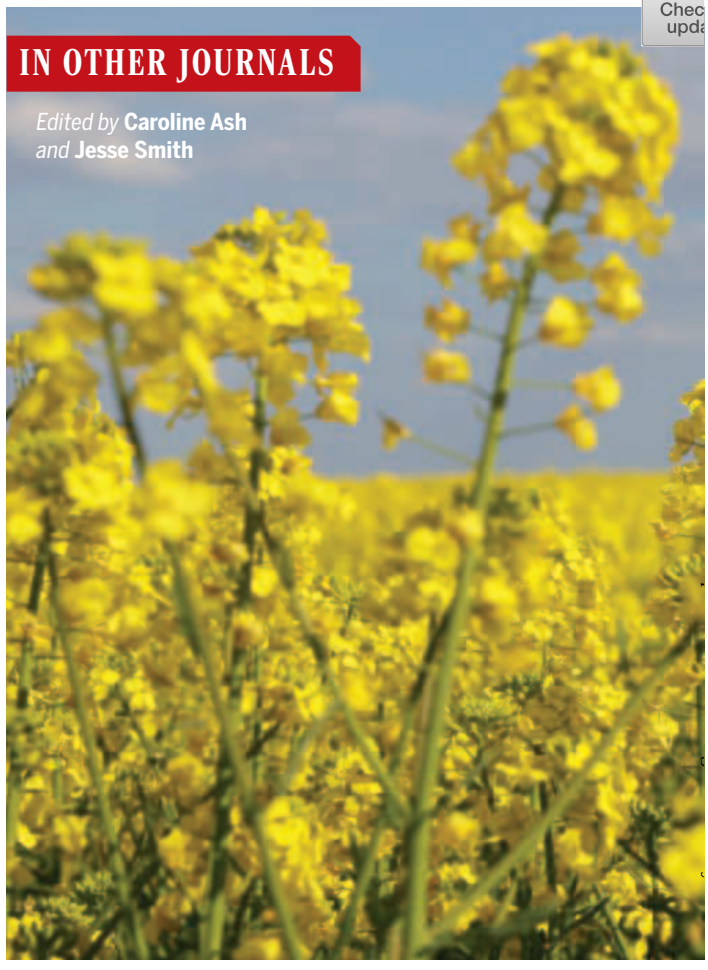
*Sci. Signal.* (2023) 10.1126/scisig.eadh0601

HIV

## bNAbs for babies

Broadly neutralizing antibodies (bNAbs) have shown promise as a companion or alternative to antiretroviral therapy (ART) for HIV-1, although their efficacy in the pediatric population remains unclear. Shapiro *et al.* report the results of a prospective clinical trial in which children on ART from birth in Botswana were treated with two HIV-1 bNAbs. The treatment was administered initially in combination with ART and was then given alone. Eleven of 25 infants maintained viral suppression during the bNAb-only step, and the authors found that these infants had more favorable HIV-1 reservoir characteristics, including a smaller initial proviral reservoir and susceptibility of those proviruses to bNAb neutralization. These results highlight the potential of bNAb treatment for infants and children living with HIV-1. —CSM

*Sci. Transl. Med.* (2023) 10.1126/scitranslmed.adh0004




CONSERVATION

## The human predator

Predators are formative components of ecosystems. Although modern humans tend not to think of ourselves as predators, ecologically, we are the predator with the broadest and largest impact across species. Darimont *et al.* surveyed more than 45,000 vertebrate species and estimated that over a third are "preyed" upon by humans through hunting and the pet trade, among other activities. Removal of these species by humans may have ecosystem-level impacts, as 40% of harvested species are threatened by our exploitative consumption. —SNV

*Commun. Biol.* (2023) 10.1038/s42003-023-04940-w

PREDATION

## Lesser fleas ad infinitum

We understand strategic aspects of the binary lifestyles of bacteria chiefly from model species. However, bacteria fall prey to smaller bacteria, the growth and division of which may be regulated in different ways. Using single-cell methods, Santin *et al.* systematically explored the life cycle of the bacterial predator *Bdellovibrio bacteriovorus*. This predator forms nonbinary replicative filaments within compartments in bacterial prey cells, which generate variable numbers of daughter cells. Larger prey cells mean more offspring than smaller cells, but it is the nutritional quality of the prey (experimentally manipulated through the growth medium) that ultimately determines the growth rate of the predator. That the timing of key cellular processes scales with the size and quality of prey cells indicates that this tiny predator can adjust its life cycle to the variability of the prey it encounters. —CA

*Curr. Biol.* (2023) 10.1016/j.cub.2023.04.059

PHOTO: MATTEO CHINELLATO/ALAMY STOCK PHOTO

Rapeseed, a source of vegetable oil and protein, growing in fields in Germany

AGRICULTURE

## Predicting yield in the field

Most plant gene expression studies are performed in controlled laboratory conditions. However, gene expression can be very different when plants are grown in the constantly changing environments of fields. Additionally, expression patterns can vary even between genetically identical individuals. De Meyer *et al.* sequenced transcripts from the leaves of field-grown oilseed rape (canola) plants (*Brassica napus*). Machine learning was applied to sequencing information to model phenotypic characteristics such as branch numbers, seed weight, and shoot dry weight. For example, the authors found that transcriptional levels of flowering time genes measured in autumnal leaves correlated with crop yield the following year. Therefore, the transcriptomes of individual plants can be used to understand genes and processes influencing crop yield in field conditions. —MRS

*PLoS Comput. Biol.* (2023)
10.1371/journal.pcbi.1011161

SYNTHETIC BIOLOGY

## Moving toward motile artificial cells

Ideally, synthetic cells need to be motile. For movement, artificial cells would need to be equipped with complex multiprotein machineries. Using the bacterial MinD and MinE proteins, Fu *et al.* achieved persistent directional movement of cell-mimicking liposomes on supported lipid bilayers. The Min proteins use energy from ATP to self-organize asymmetrically in the liposome membrane, leading to shape changes in the "cell" and mechanical force generation that promotes movement. This series of events reorganizes the proteins into a mechanochemical feedback loop sufficient to drive continuous motion. —LBR

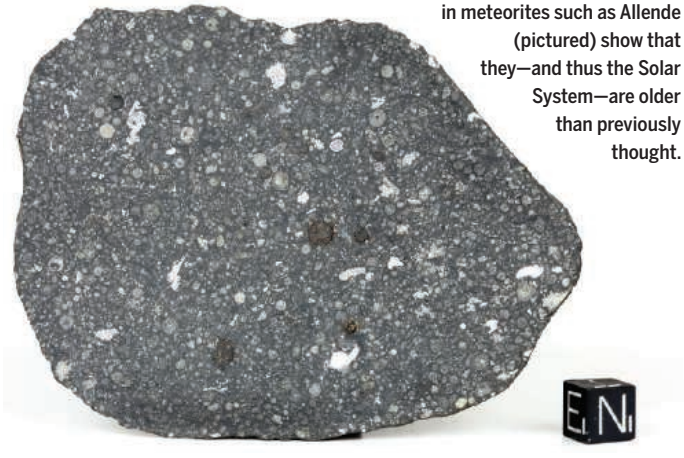
*Nat. Phys.* (2023)
10.1038/s41567-023-02058-8

METEORITICS

## Is the Solar System older than was thought?

Calcium–aluminum-rich inclusions (CAIs) in meteorites are the oldest solids that formed in the Solar System. By convention, the age of CAIs (4567.3 million years) is assumed to set the beginning of Solar System formation. Desch *et al.* reexamined this assumption using the aluminum–magnesium and lead–lead isotopic chronometers recorded in meteorites, and argue that the ages of CAIs were reset by transient heating. Extrapolating other meteorite data that were not reset indicates that the Solar System began to form 1.1 million years earlier than the apparent age of CAIs. This is a small correction compared with the age of the Solar System, but a large one when examining the relative timings of early formation events. —KTS

*Icarus* (2023)
10.1016/j.icarus.2023.115607




Revised ages for CAIs found in meteorites such as Allende (pictured) show that they—and thus the Solar System—are older than previously thought.

ORGANOMETALLICS

## Sodium nucleophiles react swimmingly

For decades, lithium has been the element of choice to stabilize negatively charged carbon in chemical reactions. Organosodium compounds, by contrast, have been plagued by solubility problems. Dilauro *et al.* report that the key all along was just to let the sodium compounds react in suspensions. Specifically, they treated organohalides in hexane with solid sodium to generate carbanions and then introduced these reagents into suspensions of electrophiles in water. The ensuing reactions proceeded cleanly and rapidly to produce coupling products, even at gram scale. —JSY

*Angew. Chem. Int. Ed.* (2023)
10.1002/anie.202304720

PROTEIN THERAPEUTICS

## Drugs with a switch

Controlling the activity of therapeutics separately from their clearance from the body would add an additional layer of control and safety, especially for protein-based drugs that activate immune responses. Marchand *et al.* inserted a previously developed OFF switch between the two domains of an antibody and performed additional engineering to optimize the binding affinity for a competing ligand. Adding this competing molecule disrupted the protein complex and led to its degradation. The authors show that this approach could also be used to generate switchable fusions with cytokines or other therapeutic proteins. —MAF

*ACS Chem. Biol.* (2023)
10.1021/acschembio.3c00012

RESEARCH ARTICLE SUMMARY

FOREST PHENOLOGY

# Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice

**Constantin M. Zohner*, Leila Mirzagholi, Susanne S. Renner, Lidong Mo, Dominic Rebindaine, Raymo Bucher, Daniel Palouš, Yann Vitasse, Yongshuo H. Fu, Benjamin D. Stocker, Thomas W. Crowther**

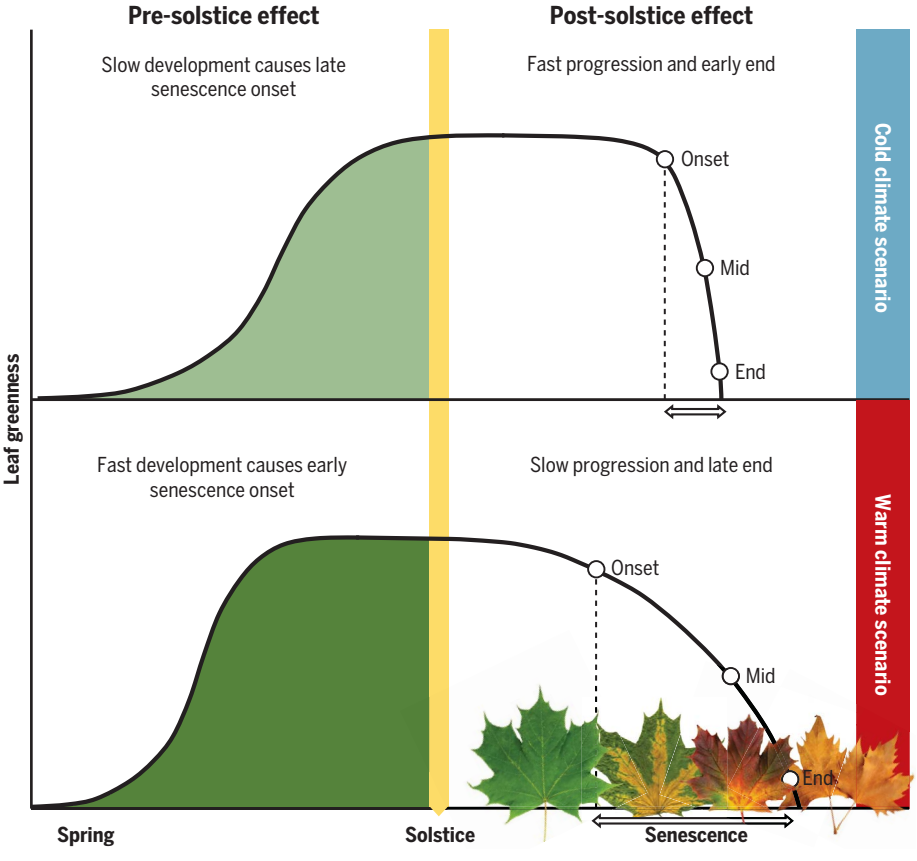
**INTRODUCTION:** Ongoing climate change is causing rapid shifts in plant phenology, with far-reaching effects on the terrestrial carbon cycle and biodiversity. While advances in spring leaf-out dates in temperate and boreal forests are well documented, the effects on autumn leaf senescence are less clear. This is because leaf senescence is not only affected by temperature but also by day length and vegetation activity early in the season in ways that are poorly understood. Accurately predicting future growing-season lengths and plant photosynthesis requires a better understanding of these interacting mechanisms at broad spatial scales.

**RATIONALE:** Local observations and experiments suggest that early-season warming, causing earlier spring leaf-out and faster plant development, tends to advance autumn senescence dates. Conversely, late-season warming tends to delay autumn senescence. If true more generally, then climate warming has opposing effects at the start and end of the growing season, with a reversal of effects somewhere in between. To test the generality of the opposing effects of climate warming on leaf senescence in Northern Hemisphere forests, we used satellite, ground, and carbon flux data, as well as controlled experiments.

**RESULTS:** Our results revealed that warming early and late in the growing season indeed has contrasting effects on leaf senescence, with a reversal occurring after the summer solstice. Across 84% of the northern forest area, we found that warmer temperatures and increased vegetation activity before the solstice advance the onset of senescence by 1.9 ± 0.1 days per °C, whereas warmer post-solstice temperatures slow the progression of senescence by 2.6 ± 0.1 days per °C. Between 1966 and 2015, the earlier onset of senescence has led to advances of 0.20 ± 0.07 days per year of the date at which autumn temperature starts to drive senescence progression. By contrast, mid-senescence continues to occur slightly later by 0.04 ± 0.01 days per year, leading to a lengthening of the autumnal senescence period.

In our experiments, warmer pre-solstice temperatures also led to earlier primary growth cessation (bud set), demonstrating that the impact of a warmer pre-solstice period extends beyond leaf development and life span. This highlights the crucial role of overall plant development and sink activity before the summer solstice in determining growing-season length.

**CONCLUSION:** We have developed a unified explanatory framework for predicting autumn phenology, showing that leaf senescence now tends to begin earlier, because of increased pre-solstice vegetation activity, but progresses more slowly, such that the end of senescence occurs later. The reversal in trees' responsiveness to warming after the summer solstice likely is triggered by changes in day length and allows them to initiate the physiological processes of leaf senescence and nutrient resorption in a fine-tuned balance between source and sink dynamics. Our results demonstrate the impact of developmental constraints (from cell and tissue growth) on autumn leaf senescence and forest productivity, affecting trends in growing-season length across the entire Northern Hemisphere. These insights provide a better understanding of the changes in growing seasons and carbon uptake of temperate and boreal forests under climate change. ■




**Autumn phenological responses to pre-solstice and post-solstice climate warming.** In cold years, slow development before the summer solstice delays the onset of senescence, and cold autumn temperatures accelerate senescence progression. In warm years, fast development before the summer solstice advances senescence onset, and warm autumn temperatures slow senescence progression, delaying the end of senescence.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: constantin.zohner@gmail.com
Cite this article as C. M. Zohner *et al.*, *Science* **381**, eadf5098 (2023). DOI: 10.1126/science.adf5098

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adf5098

RESEARCH ARTICLE

FOREST PHENOLOGY

# Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice

Constantin M. Zohner[1]*, Leila Mirzagholi[1,2], Susanne S. Renner[3], Lidong Mo[1], Dominic Rebindaine[1], Raymo Bucher[1], Daniel Palouš[1,4], Yann Vitasse[5], Yongshuo H. Fu[6], Benjamin D. Stocker[7,8], Thomas W. Crowther[1]

Climate change is shifting the growing seasons of plants, affecting species performance and biogeochemical cycles. Yet how the timing of autumn leaf senescence in Northern Hemisphere forests will change remains uncertain. Using satellite, ground, carbon flux, and experimental data, we show that early-season and late-season warming have opposite effects on leaf senescence, with a reversal occurring after the year's longest day (the summer solstice). Across 84% of the northern forest area, increased temperature and vegetation activity before the solstice led to an earlier senescence onset of, on average, 1.9 ± 0.1 days per °C, whereas warmer post-solstice temperatures extended senescence duration by 2.6 ± 0.1 days per °C. The current trajectories toward an earlier onset and slowed progression of senescence affect Northern Hemisphere–wide trends in growing-season length and forest productivity.

The phenological cycles of trees exert a strong control on the structure and functioning of ecosystems (*1*, *2*) and global carbon, water, and nutrient cycles (*3*–*5*). Anthropogenic climate change has resulted in shifts in the growing seasons of temperate and boreal trees, with the start of the season today occurring, on average, two weeks earlier than it did during the 19th and 20th centuries (*6*) and the end of the season (EOS) being delayed (*4*, *7*, *8*). Each day of a longer growing season may increase net ecosystem carbon uptake by 3.0 to 9.8 gC $m^{-2}$ (*4*). Yet, owing to the complex and interacting effects of growing-season climate and the annual day-length cycle, the direction of EOS changes in response to climate change and the cascading effects on ecosystem productivity remain highly uncertain (*9*–*13*).

Characterizing the interplay among the environmental drivers of EOS at broad spatial scales is integral to improving our understanding of growing-season length and tree growth. Cell division, tissue formation, and growth in northern trees are highest at the beginning of the season and decline with shortening days (*14*–*18*), the adaptive reason being the limited time remaining for tissue maturation and bud set before the first frost (*19*). Local observations and experiments have shown that early-season warming, causing earlier spring leaf-out and faster growth and tissue maturation, tends to advance EOS dates (*9*, *11*, *20*, *21*), whereas late-season warming has the opposite effect, delaying the EOS (*22*–*24*). Increased temperatures and physiological activity in the beginning of the season might drive earlier autumn senescence through a variety of possible mechanisms, including developmental and nutrient constraints (*9*, *25*, *26*), seasonal buildup of water stress (*27*, *28*), and radiation-induced leaf aging (*29*). In contrast, later in the season, a direct effect of temperature (cooling) is likely to drive the timing of autumn senescence (*10*, *24*, *30*). If these trends are correct, then climate warming has opposing effects at the start and end of the growing season, with a reversal of effects somewhere in between.

In this study, we tested whether early-season temperature and vegetation activity drive an earlier EOS across temperate and boreal forests, with day length providing the linkage between seasonal activity, air warming, and autumn phenology (Fig. 1). We expected that early-season plant development and growth determine the onset of senescence (Fig. 1, scenario 1 versus 2), while autumn temperature affects senescence progression (toward full dormancy), with faster chlorophyll breakdown in colder autumns than in warmer

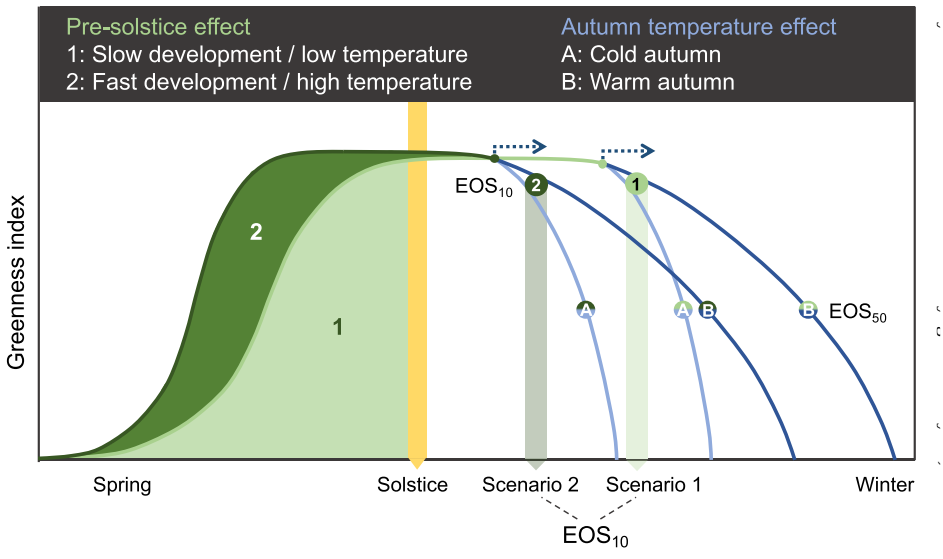



**Fig. 1. Conceptual model of autumn phenological responses to pre-solstice and post-solstice growth and temperature (solstice-as-phenology-switch hypothesis).** The onset of autumnal senescence was estimated in this study as the date when the greenness index last dropped by >10% of the seasonal maximum ($EOS_{10}$). In northern forests, stem growth, development rates, and photosynthetic capacity are highest before the summer solstice and decline with shortening days (*14*–*16*). Interannual variation in $EOS_{10}$ is a function of pre-solstice growth and development, with later $EOS_{10}$ in years with slow development and low temperature before the solstice (scenario 1) and earlier $EOS_{10}$ in years with fast development and high temperature (scenario 2). The progression of leaf senescence varies with autumn temperature, with faster chlorophyll breakdown in cold-autumn years (scenario A) than in warm-autumn years (scenario B). The dates of 50% chlorophyll loss ($EOS_{50}$) are therefore driven by pre- and post-solstice effects, whereas $EOS_{10}$ dates are mainly driven by pre-solstice effects. An earlier start of senescence in high-activity years (scenario 2) also predicts that trees will become sensitive to autumn cooling earlier than in low-activity years (see blue arrows).

[1]Institute of Integrative Biology, ETH Zurich (Swiss Federal Institute of Technology), 8092 Zurich, Switzerland. [2]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Department of Biology, Washington University in St. Louis, Saint Louis, MO 63130, USA. [4]Department of Experimental Plant Biology, Charles University in Prague, CZ 128 44 Prague, Czech Republic. [5]WSL Swiss Federal Institute for Forest, Snow and Landscape Research, 8903 Birmensdorf, Switzerland. [6]College of Water Sciences, Beijing Normal University, Beijing 100875, China. [7]Institute of Geography, University of Bern, 3012 Bern, Switzerland. [8]Oeschger Centre for Climate Change Research, University of Bern, 3012 Bern, Switzerland.
*Corresponding author. Email: constantin.zohner@gmail.com

autumns (*30*, *31*) (Fig. 1, scenario A versus B). Continued acceleration of early-season growth and development under climate warming (*4*, *5*, *32*, *33*) might thus cause an ever earlier EOS onset, whereas the progression of senescence should be slowed down by warmer autumns (*22*, *23*), lengthening the overall senescence period. This leads to four fundamental predictions that were tested in this study: (i) Enhanced pre-solstice temperature and vegetation activity drives an earlier senescence onset (Fig. 1, scenario 1 versus 2). (ii) Growth and temperature effects on senescence dates reverse around the time of the summer solstice. (iii) Autumn temperature affects the speed of senescence, delaying its later stages (scenario A versus B), but has little effect on its start. (iv) The date when trees become sensitive to autumn cooling (blue arrows in Fig. 1) has advanced over recent decades because of an earlier onset of senescence.

To test these hypotheses, we combined phenology data from satellite observations across Northern Hemisphere temperate and boreal forests (*34*), ground observations from European (*35*) and American (*36*) deciduous trees, eddy covariance flux tower measurements (*37*), and controlled experiments on European beech (*38*). Seasonal vegetation activity was estimated using direct measurements of leaf-level gas exchange and ecosystem carbon fluxes (*37*), as well as photosynthesis models [satellite-derived gross primary productivity (GPP) (*39*) and the Lund–Potsdam–Jena General Ecosystem Simulator (LPJ-GUESS) model (*9*)]. We then ran linear models to test for the monthly and seasonal effects of photosynthesis, temperature, shortwave radiation, and water availability on EOS dates. The satellite data allowed us to differentiate between the onset of senescence and its progression by analyzing the dates when greenness had dropped by 10% ($EOS_{10}$) or 50% ($EOS_{50}$) relative to the seasonal maximum. The experiments allowed us to directly test for seasonal variation in the effects of air temperature and radiation. Finally, we mapped the relative effects of early-season vegetation activity and late-season climate across Northern Hemisphere forests to test for possible biogeographic patterns in the drivers of autumn senescence.

### Effect of temperature and vegetation activity on senescence dates reverses after the summer solstice

Satellite-based phenology data (Figs. 2 and 3 and figs. S1 to S3), European (Fig. 4) and American (fig. S4) ground observations, flux tower measurements (figs. S5 and S6), and experiments (fig. S7) all revealed a consistent advancing effect of pre-solstice (i.e., before 21 June) temperature and productivity on EOS dates, which declined after the summer solstice. Thus, across 84% of the northern forest area [18% (Fig. 3C) and 22% (fig. S1C) significant at $P < 0.05$], increased pre-solstice daytime temperature ($T_{day}$) and photosynthesis led to an earlier onset of senescence: each 1°C increase in pre-solstice temperature resulted, on average, in 1.9 ± 0.1 (mean ± 2 SE) days earlier $EOS_{10}$ (Fig. 3D). Similarly, each

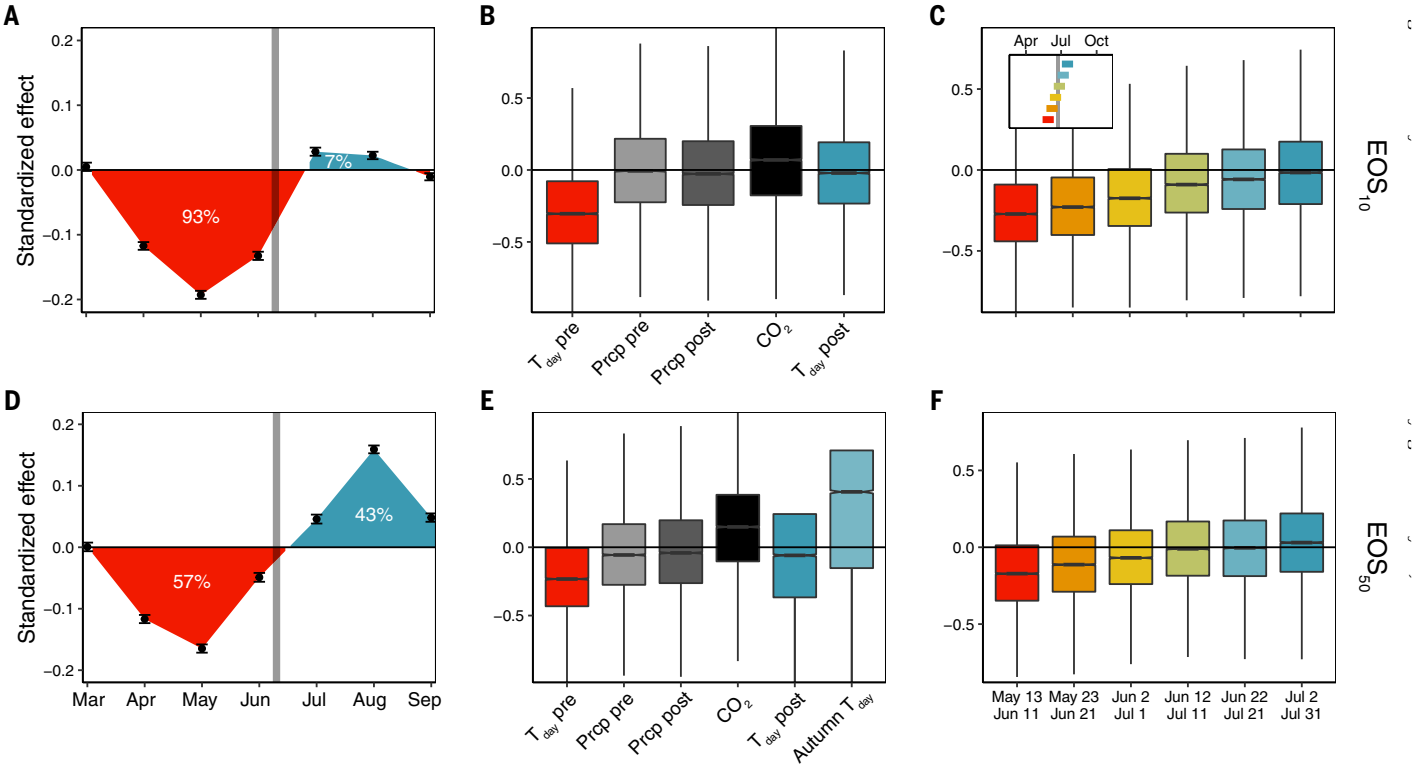



**Fig. 2. Effect of temperature on the timing of autumn senescence in northern forests reverses after the summer solstice.** Autumn phenology is represented as $EOS_{10}$ (A to C) or $EOS_{50}$ dates (D to F) from satellite observations (see Fig. 3A). (**A** and **D**) Relationship between mean monthly $T_{day}$ and EOS dates (means ± 95% confidence ranges) from multiple linear regression models. Percentages reflect the total positive and negative areas under the curve, meaning the relative advancing versus delaying effects of seasonal $T_{day}$. (**B** and **E**) Effects of pre-solstice daytime temperature ($T_{day}$ pre; 20 March to solstice) and precipitation (Prcp pre), post-solstice daytime temperature ($T_{day}$ post; solstice to mean pixel-level EOS date) and precipitation (Prcp post), and atmospheric $CO_2$ on EOS dates [(E) additionally includes the effect of autumn $T_{day}$]. (**C** and **F**) The univariate effects of 1-month-long $T_{day}$ intervals around the summer solstice [13 May to 11 June, 23 May to 21 June, 2 June to 1 July, 12 June to 11 July, 22 June to 21 July, and 2 July to 31 July; the inset in (C) illustrates the time of year of each period, and the gray vertical line represents the summer solstice] on EOS dates. Analyses show relative effect sizes from pixel-level linear models. To examine relative effect sizes, both predictor and dependent variables were standardized by subtracting their mean and dividing by one standard deviation before analysis.

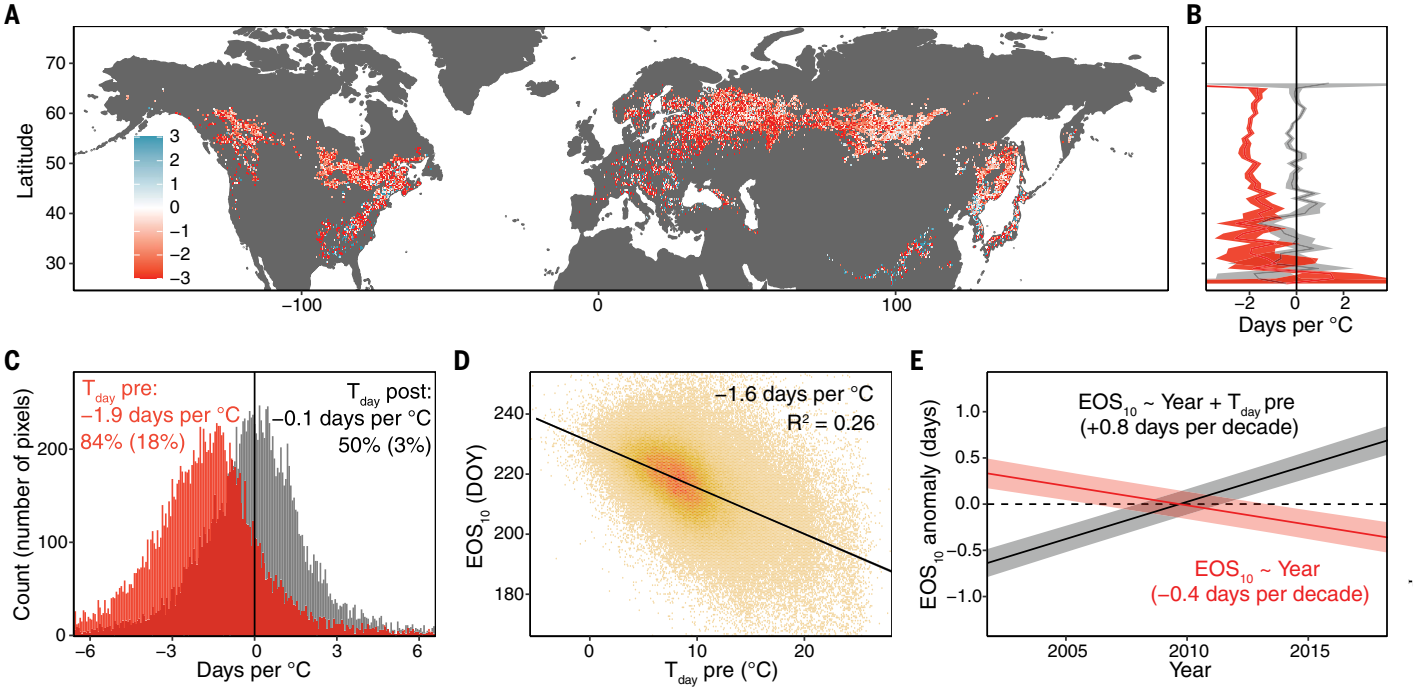



**Fig. 3. Satellite observations reveal consistent advances in the onset of senescence ($EOS_{10}$) across northern forests in response to increased pre-solstice temperature.** (**A**) Map showing the effects of pre-solstice (20 March to solstice) $T_{day}$ on $EOS_{10}$ timing at 0.25° resolution from linear models, including pre-solstice $T_{day}$ and post-solstice (solstice to mean pixel-level EOS date) $T_{day}$ as predictor variables. Red pixels indicate an earlier $EOS_{10}$ under increased pre-solstice $T_{day}$, blue pixels indicate a delayed $EOS_{10}$. (**B**) Regression coefficient means and 95% confidence ranges summarized for each degree latitude (pre-solstice effects in red, post-solstice effects in black). (**C**) The distribution of the pre-solstice and post-solstice $T_{day}$ effects across all pixels. Mean pre- and post-solstice $T_{day}$ regression coefficients and the percentage of pixels with an advancing pre-solstice $T_{day}$ effect or delaying post-solstice $T_{day}$ effect shown as red and black text, respectively (percentage of significant pixels at $P < 0.05$ in brackets). (**D**) Partial effect of pre-solstice $T_{day}$ on $EOS_{10}$ dates, controlling for year (fixed effect) and site (random effect). (**E**) Temporal trends in $EOS_{10}$ anomalies (means ± 95% confidence ranges) from mixed-effects models with year as a single fixed effect (red line and equation) or including pre-solstice $T_{day}$ as an additional fixed effect (black line and equation). Pixels were treated as grouping variables of random intercepts.

10% increase in pre-solstice photosynthesis resulted in 3.6 ± 0.1 days earlier $EOS_{10}$ (satellite data; fig. S1) or 3.7 ± 1.2 days earlier $PD_{25}$ (date when photosynthesis had dropped by 25% relative to the seasonal maximum according to flux tower data; fig. S6). A significant delaying effect of pre-solstice temperature and photosynthesis was found for <1% of northern forest pixels.

Post-solstice temperature had a small effect on the onset of senescence (see Figs. 2, A and B, and 5A for $EOS_{10}$ dates, and fig. S8, A and B, for $EOS_{start}$ dates as an alternative measure of senescence onset) but prolonged its duration (number of days from $EOS_{10}$ to $EOS_{85}$, $EOS_{10}$ to $EOS_{50}$, or $EOS_{50}$ to $EOS_{85}$) by 1.3 to 3.5 days per °C (fig. S9), explaining the large effect of autumn temperatures on $EOS_{50}$ dates (Fig. 2, D and E, and fig. S2, C to E, for satellite observations and Figs. 4 and 5C for European plots). Precipitation and $CO_2$ levels had comparatively small effects (Fig. 2, B and E, and fig. S10).

The reversal of the effects of air temperature and vegetation activity on EOS dates after the summer solstice was consistent across (i) all EOS metrics used here, that is, the onset of senescence ($EOS_{10}$ or $EOS_{start}$; Fig. 2, A to C, and fig. S8) and mid-senescence ($EOS_{50}$; Fig. 2, D to F); (ii) continents and forest types (fig. S11); and (iii) a set of alternative variables linked to growing-season activity and development, namely $T_{day}$ (Fig. 3 and fig. S2) and photosynthetic activity (figs. S1 and S3). Along the full latitudinal gradient (30°N to 65°N) studied here, the period during which vegetation productivity had an advancing effect on $EOS_{10}$ dates consistently ended after the solstice, at ~26 June (fig. S12). The effect reversal after the summer solstice was further supported by an analysis that used 10-day moving steps around the solstice (Fig. 2, C and F). To further test for the importance of separating pre- and post-solstice variables for EOS predictions, we ran leave-one-out cross-validation of models that included (i) only pre-solstice, (ii) only post-solstice, or (iii) both types of variables as predictors of EOS dates (*38*). In full agreement with our hypothesis (Fig. 1), the onset of senescence ($EOS_{10}$) was better explained by pre-solstice than by post-solstice variables, and the full model—including both pre-and post-solstice variables—showed only slightly better performance than the pre-solstice model (fig. S13, A and D). Because autumn temperature modulates the progression of senescence, $EOS_{50}$ dates were slightly better explained by post-solstice than by pre-solstice variables, with the combination of both variables yielding the best predictions (fig. S13, B, C, E, and F).

In line with the satellite observations, high pre-solstice temperature and productivity correlated with advanced $EOS_{50}$ dates in the European and American plot data, across all species (Fig. 4B and fig. S4) and across a set of alternative variables [$T_{day}$ (Fig. 4) and LPJ-GUESS model–derived photosynthesis (fig. S15)]. On the basis of these findings, we ran multivariate mixed models for the European plot data, including pre-solstice and post-solstice (solstice to mean $EOS_{50}$) temperature or photosynthesis and precipitation, $CO_2$ levels, and autumn nighttime temperature (autumn $T_{night}$) to determine their relative importance. Pre-solstice temperature (or photosynthesis) had the strongest advancing effect on $EOS_{50}$ dates, whereas autumn $T_{night}$ had the strongest

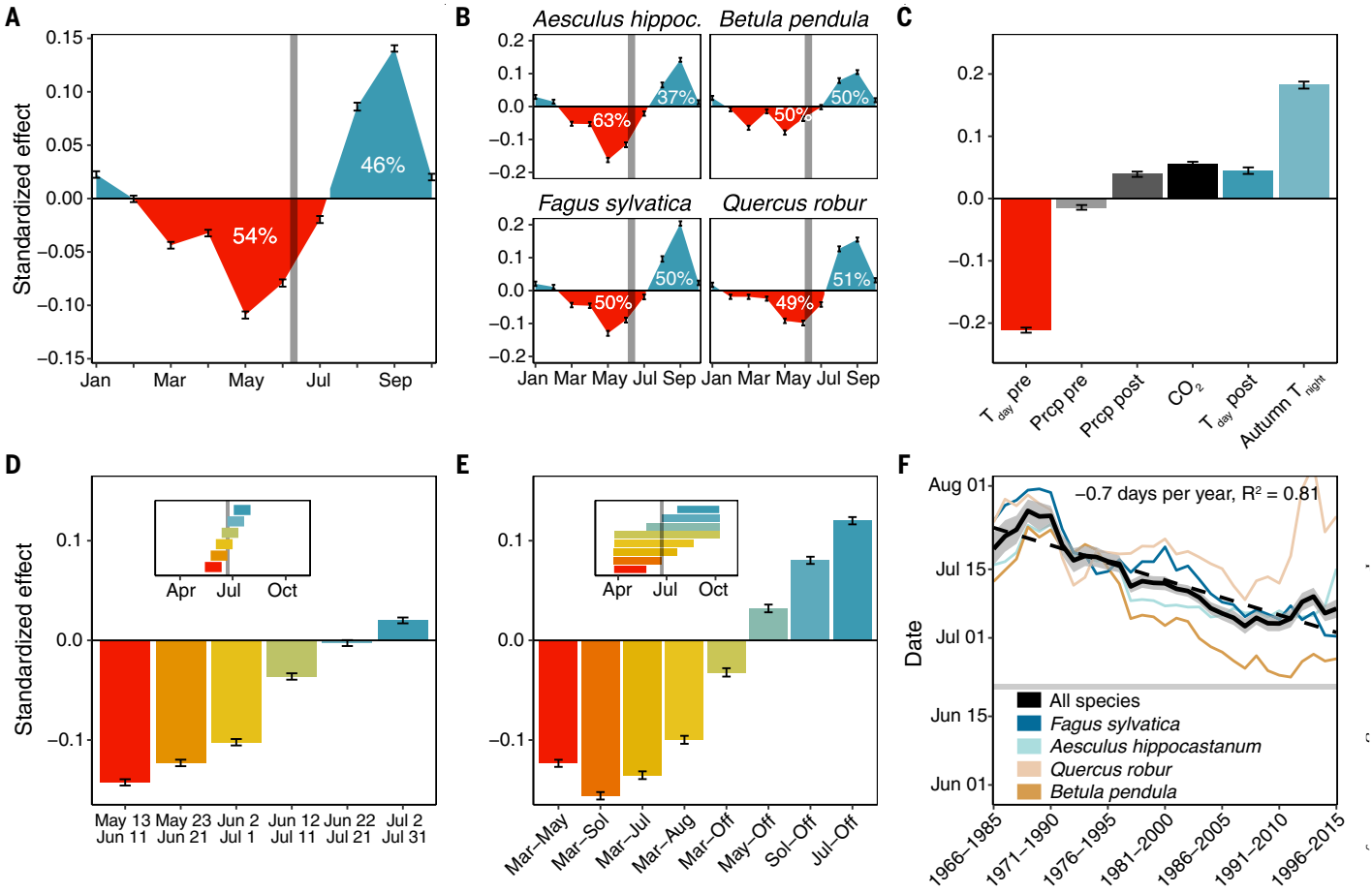



**Fig. 4. The seasonal effects of temperature on interannual variation in mid-senescence ($EOS_{50}$ dates) from European long-term observations (PEP725 data).** (**A**) Effects of monthly (January to October) $T_{day}$ on autumn senescence dates. Percentages reflect the total positive and negative areas under the curve, meaning the relative advancing versus delaying effects of temperature. (**B**) Species-level results. (**C**) The relative effects of pre-solstice (20 March to solstice) and post-solstice (solstice to mean $EOS_{50}$) temperature, pre-solstice and post-solstice precipitation, atmospheric $CO_2$, and autumn nighttime temperature (Autumn $T_{night}$). (**D**) Effects of 1-month-long temperature intervals around the summer solstice (13 May to 11 June, 23 May to 21 June, 2 June to 1 July, 12 June to 11 July, 22 June to 21 July, and 2 July to 31 July; see inset), including the respective temperature interval as a single fixed effect. (**E**) The effects of mean temperature from 20 March to 22 May (Mar–May), 20 March to solstice (Mar–Sol), 20 March to 21 July (Mar–Jul), 20 March to 20 August (Mar–Aug), 20 March to mean $EOS_{50}$ (Mar–Off), 22 May to mean $EOS_{50}$ (May–Off), solstice to mean $EOS_{50}$ (Sol–Off), and 21 July to mean $EOS_{50}$ (Jul–Off), including the respective temperature interval as fixed effect. (**F**) Moving-window analysis, showing the "reversal" dates when the temperature effect switches from negative to positive for each 20-year time period from 1966 to 2015 [based on monthly correlations; see (A) to (C)]. On average, the reversal date advanced by 0.7 days per year. Analyses show effect size means ± 2 SE from linear mixed models, including time series and species [(A) and (C) to (E)] or only time series [(B) and (F)] as random effects, with both predictor and dependent variables standardized. $R^2$, coefficient of determination.

delaying effect on $EOS_{50}$ dates, with both effects being more than three times greater than that of precipitation and atmospheric $CO_2$ (Fig. 4C). EOS predictions showed that the full model representing both pre- and post-solstice effects captured within-site $EOS_{50}$ trends in response to mean annual temperature (advance of 0.4 ± 0.1 days per °C increase in mean annual temperature; fig. S16). In contrast, a model representing only post-solstice temperature and precipitation predicted delays of 0.7 ± 0.1 days per °C, whereas a pre-solstice model predicted advances of 0.9 ± 0.1 days per °C. These findings demonstrate that information on both pre- and post-solstice climate is necessary to reproduce the observed $EOS_{50}$ responses to rising temperature.

### Ongoing trends toward an earlier start, slowed progression, and later end of senescence

Our finding that the onset of senescence is driven by pre-solstice vegetation activity and development, while the speed of its progression depends on autumn temperature (fig. S9), suggests that global warming leads to an earlier start and slowed progression of senescence (scenario 2B in Fig. 1). Indeed, across all analyzed northern forest pixels, the onset of senescence ($EOS_{10}$ date) has advanced by an average of −0.4 ± 0.1 days per decade between 2001 and 2018 (Fig. 3E), in parallel with increased pre-solstice vegetation productivity (fig. S14, A and B), with the strongest advances in $EOS_{10}$ dates found for regions with the largest increase in pre-solstice productivity (fig. S14K). When removing the effect of pre-solstice temperature or photosynthesis on senescence onset by including either variable as a fixed effect in addition to year, the model instead predicted delays in $EOS_{10}$ dates of 0.8 to 1.4 days per decade (Fig. 3E and fig. S1E), underscoring how strongly temperature-driven pre-solstice vegetation activity affects the start of senescence. The warmer autumns (fig. S14, I and J) are instead slowing the progression of

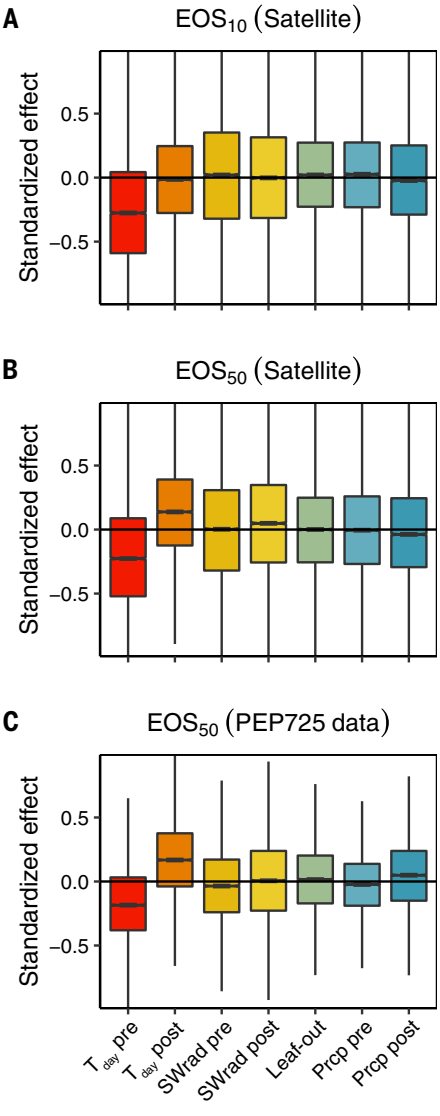



**Fig. 5. The effects of temperature, radiation, spring leaf-out dates, and precipitation on interannual variation in the timing of $EOS_{10}$ and $EOS_{50}$.** Panels show results for satellite-derived $EOS_{10}$ dates (**A**), satellite-derived $EOS_{50}$ dates (**B**), and ground-sourced $EOS_{50}$ dates (**C**). All linear models include mean $T_{day}$ and shortwave radiation (SWrad) and the sums of precipitation (prcp) from 20 March to 21 June (pre-solstice) and from 22 June to the mean EOS date of each time series (post-solstice), as well as spring leaf-out dates, as predictor variables. Models were run at the pixel level [(A) and (B)] or individual level (C), and boxplots show the spread of effect sizes. All variables were standardized to allow for direct effect size comparison.

senescence, leading to delays in $EOS_{50}$ dates (fig. S2G), especially in regions with the largest increases in autumn temperature (Fig. S14M).

The offsetting effect of pre-solstice vegetation activity on autumn warming–induced delays in $EOS_{50}$ is also apparent in regional trends over the past 70 years (time series and species as random effects; fig. S17). On average, European $EOS_{50}$ dates have been delayed by only +0.35 ± 0.02 days per decade (fig. S17B). By contrast, when keeping pre-solstice productivity constant by including it as a fixed effect in addition to year, the model predicts a delay of +0.81 ± 0.03 days per decade (fig. S17D), showing that the increase in pre-solstice vegetation productivity has offset up to ~60% of the delay in $EOS_{50}$ that would have occurred had pre-solstice productivity not increased. This explains why $EOS_{50}$ delays have contributed only ~15% (2.4 ± 0.2 days) to the 16.7 ± 0.4 day–long extension of the growing season that has occurred over the past 70 years (fig. S17, A and B).

The longer senescence duration (fig. S9) and delayed $EOS_{50}$ dates (when greenness has dropped by 50%) under warmer autumns reveal how autumn temperature modulates senescence (Figs. 4C and 5, B and C, and fig. S18, C and D). However, if increased pre-solstice vegetation development (fig. S14A) is the main driver of an earlier onset of EOS, one should find an ever earlier susceptibility of trees to autumn cooling. We tested this by using temporal moving-window analyses on the European long-term observations and found that the effect reversal dates, at which increased temperature and productivity start to be associated with delayed $EOS_{50}$ dates, have indeed advanced, by an average of −0.7 to −1.0 days per year (Fig. 4F and fig. S15F or fig. S19B for a shorter moving window). This is also reflected in the moving windows of monthly effect sizes, which show that July photosynthesis has become more strongly associated with delayed $EOS_{50}$ dates over recent decades (fig. S19C). As an alternative method of determining when autumn cooling starts driving senescence progression, we modeled the autumn period best explaining $EOS_{50}$ dates and found that it has advanced by −0.20 ± 0.07 days per year for the 1966–2015 period (fig. S20A) or by −0.43 ± 0.09 days per year for the 1981–2015 period (fig. S20B). This earlier start of the period when trees react to autumn cooling provides further evidence for an earlier onset of senescence in response to increased early-season development.

An important consideration of our autumn senescence model presented in Fig. 1 is that the effect reversal date—the compensatory point of the advancing and delaying effects—should be flexible despite day length likely triggering the initial decline in the advancing effect after the summer solstice (*14*, *15*, *17*). This is because the actual reversal date of the antagonistic effects, meaning the period when trees become sensitive to cooling in late summer, is a function of warming and development before the summer solstice, which will vary (Fig. 1). Involvement of day length and a flexible effect reversal date are therefore not mutually exclusive ideas but rather necessary components of the same model.

### Factors governing the link between pre-solstice temperature and senescence onset

Previous research on European deciduous trees has suggested a negative feedback between growing-season productivity and autumn phenology, with increased productivity driving earlier senescence (*9*). The data analyzed here now reveal that the productivity before the summer solstice is indeed linked to earlier senescence dates and that this holds across the entire Northern Hemisphere temperate and boreal forest biome, implying a widespread constraint on future growing-season extensions in response to global warming. To disentangle the environmental drivers of this feedback, we ran multiple linear regression models that included air temperature, solar radiation, water availability, and spring leaf-out dates as predictor variables, all of which have been shown to affect leaf senescence dates (*9*, *11*, *20*, *26*, *28*, *40*) (Fig. 5 and figs. S21 or S22 using soil moisture instead of precipitation to represent water availability). Across the majority of pixels or sites, pre-solstice temperature was the main driver of $EOS_{10}$ dates (Fig. 5A), whereas $EOS_{50}$ dates were driven by both pre- and post-solstice temperatures, with opposite effect directions (Fig. 5, B and C). This suggests that temperature-driven development and growth—rather than radiation-induced leaf aging, drought, or leaf-out per se (i.e., constrained leaf life span)—are driving the advancing effect of early-season vegetation activity on autumn phenology.

To further isolate the mechanisms driving the discovered reversal of the effects of global warming around the summer solstice, we conducted an experiment on a dominant European tree species (*Fagus sylvatica*), in which we cooled (chamber temperature set to 10°C during the day and 5°C at night) and shaded (~84% light reduction) saplings during different times of the season. Pre-solstice temperature again had a strong advancing effect on autumn phenology, with cooling of trees in June causing a delay in $EOS_{10}$ and $EOS_{50}$ dates of +16.5 ± 6.6 days and +10.2 ± 2.5 days (mean ± SE), respectively. Cooling in July had no effect, and August cooling tended to advance EOS dates (figs. S7A and S23A), in full agreement with the global-scale remote sensing data and the ground observations. The effect of shading was small before the summer solstice and most pronounced during July—the month with the highest mean daily radiation and temperature—with $EOS_{50}$ delayed by +6.5 ± 2.8 days under shade conditions. Radiation effects thus followed a different seasonal pattern than temperature, supporting a direct effect of radiation on leaf aging (*26*, *29*). Summer

photosynthesis was equally reduced in both the shade and the cooling treatments by 52 to 72% compared with the control (fig. S24). That pre-solstice temperature, but not pre-solstice light availability, affected EOS dates provides further support for the idea that accelerated development and growth under warmer temperatures, rather than photosynthesis- or radiation–induced leaf aging, are the main drivers of the pre-solstice effects on senescence dates. Yet given that both plant developmental speed (tissue expansion and meristematic activity) and photosynthesis early in the season are strongly driven by temperature (*41*), both processes are linked, and pre-solstice GPP therefore appears to be a reliable proxy of this effect (fig. S1).

An important future avenue will be to disentangle the relative roles of developmental processes within leaves (*42*) versus growth processes in the rest of the plant (*25*, *41*) in driving earlier senescence under increased pre-solstice temperatures. That pre-solstice temperatures and productivity advanced senescence even in species with continuous leaf production throughout the season, such as *Betula pendula* (Fig. 4B and fig. S15B), suggests that leaf development and longevity cannot be the sole drivers, implying involvement of growth (sink) processes in other organs, including wood, roots, and buds. This idea was further supported by a second experiment in which we tested the response of primary growth cessation (bud set) to pre- and post-solstice warming. The results showed that pre-solstice warming (22 May to 21 June) advanced bud set in *F. sylvatica* by 4.4 ± 1.8 days, whereas post-solstice warming (22 June to 21 July) delayed bud set by 4.9 ± 1.8 days (fig. S25), consistent with an effect reversal after the summer solstice. The negative effect of warmer pre-solstice temperatures on growing-season duration is therefore not limited to leaf development and life span (*23*, *42*, *43*), suggesting an important role of overall plant growth and sink activity before the solstice in determining autumn growth cessation and leaf senescence (*14*, *16*, *17*, *41*).

The observed EOS advances under elevated pre-solstice temperature and growth may additionally be linked to availability of soil resources throughout the growing season. Optimal growth conditions in spring (high temperatures and sufficient precipitation) may deplete nitrogen availability to trees, which in turn may drive early bud set and senescence (*25*, *26*, *44*). This effect may be further enhanced by increased C:N ratios of leaves under rising atmospheric $CO_2$, causing microbial immobilization and reduced nitrogen accessibility (*45*, *46*).

## Conclusions

Our investigation of the seasonal drivers of autumn phenology shows a consistent reversal after the summer solstice in the effects of climate warming on leaf senescence (i) across large biogeographic ranges with varying pre-solstice growth (satellite data); (ii) in multiple, broad-ranged tree species with different spring phenologies (ground data); (iii) in eddy covariance measurements; and (iv) under controlled experimental conditions. These findings reveal important constraints on growing-season length and photosynthetic activity, whereby earlier and enhanced activity before the summer solstice drives earlier photosynthetic declines, primary growth cessation, and senescence in autumn. The seasonal control on EOS by air temperature and vegetation activity may be mediated by the annual day-length cycle (*15*, *17*) and nutrient availability (*26*, *44*). Year-to-year differences in the onset and progression of autumn senescence thus emerge as the result of a complex synchronization between trees' developmental states, seasonal variation in the circadian rhythm, and climate fluctuations. This mediation by the annual day-length cycle provides a unifying framework to explain previous results, in which the magnitude and direction of climate warming effects on autumn phenology varied (*9*, *12*, *22*, *23*, *33*, *47*–*50*), largely because studies did not disentangle pre- and post-solstice variables. It now is clear that warmer temperatures and increased vegetation activity before the summer solstice drive an earlier onset of senescence, whereas, in agreement with previous studies (*12*, *47*, *51*), warmer temperatures after the solstice slow down the progression of senescence, predicting that senescence will continue to start earlier but progress more slowly.

The reversal in how trees respond to temperature during the summer likely evolved as an adaptation to seasonal climates with harsh winters because it allows plants to reliably anticipate the progression of the seasons and prepare for winter dormancy long before the temperature starts dropping (*19*). The solstice switch in trees' physiological responsiveness to temperature calibrates their seasonal rhythms and mediates how they react to warm or cool temperatures, now and in the future. This enables trees to initiate tissue maturation and the physiological processes of leaf senescence and nutrient resorption (*15*) in a fine-tuned balance between source and sink dynamics. Improved models of plant development and growth under climate change will need to incorporate the reversal of warming effects after the summer solstice.

## Methods summary

### Satellite observations

Start-of-season (SOS) and EOS dates for Northern Hemisphere forests from 2001 to 2018 came from the MODIS Global Vegetation Phenology product (*34*). SOS was defined as the date when satellite-derived greenness had reached 15% of its annual maximum. EOS was defined as the date when greenness had decreased by 10% ($EOS_{10}$), 50% ($EOS_{50}$), or 85% ($EOS_{85}$), representing the start of senescence, mid-greendown, and dormancy, respectively. As an alternative measure of the start of senescence, we used the VIIRS/NPP Land Cover Dynamics product (*52*), which uses a breakpoint algorithm to define the onset of greendown. Information on climate came from (*53*), and GPP came from the MODIS GPP product (*39*). We included GPP, $T_{day}$, $T_{night}$, shortwave radiation, $CO_2$ levels, precipitation, soil moisture, and SOS dates as covariates in our analyses.

To investigate the seasonal effects of temperature, photosynthesis, shortwave radiation, and water availability on the timing of $EOS_{50}$ dates, we aggregated data for different periods before and after the summer solstice. For each time series, we also determined the optimal autumn interval for which temperature explains most of the variation in $EOS_{50}$ dates. We quantified the importance of each covariate in driving interannual variation in EOS dates with pixel-level linear models (Figs. 2 and 3). To approximate the end date of the period during which early-season vegetation activity had an advancing effect on the onset of senescence, we tested the correlation between different GPP periods (always beginning at the SOS date) and $EOS_{10}$ dates (fig. S12). To test for decadal-scale temporal trends in EOS dates, we ran mixed-effects models including only year or year and pre-solstice temperature (Fig. 3, D and E) or GPP (fig. S1, D and E) as fixed effects and treating pixels as grouping variables of random intercepts.

### Ground observations

Ground data on 396,411 EOS dates of four dominant tree species in central Europe between 1951 and 2015 came from the Pan European Phenology Project (*35*). Climate data for the same period were obtained from the Global Land Data Assimilation System (*53*). The $EOS_{50}$ corresponded to the date when 50% of leaves had lost their green color. We included photosynthesis, $T_{day}$, $T_{night}$, shortwave radiation, $CO_2$ levels, precipitation, soil moisture, and spring leaf-out dates as covariates in our analyses. Daily photosynthesis was derived from the LPJ-GUESS photosynthesis model (*54*). As for the satellite observations, data were aggregated for monthly and longer periods before and after the summer solstice, and we also determined the optimal autumn interval for which temperature explains most of the variation in $EOS_{50}$ dates.

The importance of these variables in driving EOS dates was then evaluated with linear mixed models, including time series (unique site and species combination) and species random effects. Our final model included fixed

effects of pre- and post-solstice $T_{day}$ (Fig. 4C) or photosynthesis (fig. S15C), as these variables emerged as the strongest drivers of EOS dates.

To test whether the relative effects of variables have been changing over the past decades, we ran the mixed models separately for each 20-year (or 15-year) time period from 1966 to 2015 (or 1980 to 2015). To estimate the date at which the effect of temperature and photosynthesis reverses, we conducted moving-window analyses of the monthly temperature and photosynthesis effects. Additionally, we estimated the day at which autumn temperature starts driving senescence progression by calculating the autumn period for which temperature best explained variation in $EOS_{50}$ dates for each moving window (fig. S20). We additionally analyzed direct observations of $EOS_{50}$ dates for 10 North American tree species from 1991 to 2019 from the Phenology of Woody Species at Harvard Forest dataset (*36*) (fig. S4).

### *Flux tower measurements*

Eddy covariance data from 10 temperate forest (broadleaf deciduous/mixed) sites with >9 years of data were gathered from the FLUXNET2015: CC-BY-4.0 dataset (*37*). To represent the onset dates of photosynthesis declines (PD) in late summer, we extracted the day of year on which GPP last dropped below 10% ($PD_{10}$; fig. S5) or 25% ($PD_{25}$; fig. S6) of maximum annual GPP. We then tested for the relationships between monthly (April to September) GPP estimates and $PD_{10}$ or $PD_{25}$ dates by running mixed-effects models, including site as a random effect.

### *Experiments*

In the first experiment, we tested for the effects of seasonal variation in temperature and light on EOS dates. We exposed 3-year-old *F. sylvatica* trees to cooling or shading conditions during four 1-month-long periods (from May to August). In the cooling treatments, trees experienced a nighttime temperature of 5°C and a daytime temperature of 10°C. In the shading treatments, trees were exposed to an ~84% reduction in incoming radiation. Individual leaf senescence dates were defined as the day of year when chlorophyll content last dropped by 10% ($EOS_{10}$) or 50% ($EOS_{50}$) of the maximum chlorophyll content in summer. Leaf net photosynthesis was measured with a portable infrared gas analyzer (LI-6800, LI-COR Biosciences, Lincoln, NE). To test for differences in leaf senescence dates among treatments, we ran multivariate linear models including temperature and shade treatments as categorical variables.

In the second experiment, we tested the effects of pre- and post-solstice warming on bud set dates by exposing 4-year-old *F. sylvatica* trees to cooling conditions (8°C during day and night) during 1-month-long periods before and after the solstice. Bud set was defined as the date when the buds of an individual had reached >90% of their final length (*23*), indicating aboveground primary growth cessation (*23*, *55*). To test for differences in bud set dates and autumn bud growth rates among treatments, we ran mixed-effects models including treatment as categorical fixed effect and bud types (apical versus lateral) as a random effect.

## REFERENCES AND NOTES


1. I. Chuine, Why does phenology drive species distribution? *Philos. Trans. R. Soc. London Ser. B* **365**, 3149–3160 (2010). doi: 10.1098/rstb.2010.0142; pmid: 20819809
2. S. S. Renner, C. M. Zohner, Climate change and phenological mismatch in trophic interactions among plants, insects, and vertebrates. *Annu. Rev. Ecol. Evol. Syst.* **49**, 165–182 (2018). doi: 10.1146/annurev-ecolsys-110617-062535
3. A. D. Richardson *et al*., Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agric. For. Meteorol.* **169**, 156–173 (2013). doi: 10.1016/j.agrformet.2012.09.012
4. T. F. Keenan *et al*., Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nat. Clim. Chang.* **4**, 598–604 (2014). doi: 10.1038/nclimate2253
5. C. M. Zohner, L. Mo, T. A. M. Pugh, J. F. Bastin, T. W. Crowther, Interactive climate factors restrict future increases in spring productivity of temperate and boreal trees. *Glob. Change Biol.* **26**, 4042–4055 (2020). doi: 10.1111/gcb.15098; pmid: 32347650
6. Y. Vitasse *et al*., The great acceleration of plant phenological shifts. *Nat. Clim. Chang.* **12**, 300–302 (2022). doi: 10.1038/s41558-022-01283-y
7. N. Delpierre *et al*., Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. *Agric. For. Meteorol.* **149**, 938–948 (2009). doi: 10.1016/j.agrformet.2008.11.014
8. A. Menzel, P. Fabian, Growing season extended in Europe. *Nature* **397**, 659 (1999). doi: 10.1038/17709
9. D. Zani, T. W. Crowther, L. Mo, S. S. Renner, C. M. Zohner, Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees. *Science* **370**, 1066–1071 (2020). doi: 10.1126/science.abd8911; pmid: 33243884
10. L. Chen *et al*., Leaf senescence exhibits stronger climatic responses during warm than during cold autumns. *Nat. Clim. Chang.* **10**, 777–780 (2020). doi: 10.1038/s41558-020-0820-2
11. T. F. Keenan, A. D. Richardson, The timing of autumn senescence is affected by the timing of spring phenology: Implications for predictive models. *Glob. Change Biol.* **21**, 2634–2641 (2015). doi: 10.1111/gcb.12890; pmid: 25662890
12. G. Liu, X. Chen, Q. Zhang, W. Lang, N. Delpierre, Antagonistic effects of growing season and autumn temperatures on the timing of leaf coloration in winter deciduous trees. *Glob. Change Biol.* **24**, 3537–3545 (2018). doi: 10.1111/gcb.14095; pmid: 29460318
13. S. Piao *et al*., Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature* **451**, 49–52 (2008). doi: 10.1038/nature06444; pmid: 18172494
14. S. Rossi *et al*., Conifers in cold environments synchronize maximum growth rate of tree-ring formation with day length. *New Phytol.* **170**, 301–310 (2006). doi: 10.1111/j.1469-8137.2006.01660.x; pmid: 16608455
15. W. L. Bauerle *et al*., Photoperiodic regulation of the seasonal pattern of photosynthetic capacity and the implications for carbon cycling. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8612–8617 (2012). doi: 10.1073/pnas.1119131109; pmid: 22586103
16. S. Etzold *et al*., Number of growth days and not length of the growth period determines radial stem growth of temperate trees. *Ecol. Lett.* **25**, 427–439 (2022). doi: 10.1111/ele.13933; pmid: 34882952
17. T. Luo *et al*., Summer solstice marks a seasonal shift in temperature sensitivity of stem growth and nitrogen-use efficiency in cold-limited forests. *Agric. For. Meteorol.* **248**, 469–478 (2018). doi: 10.1016/j.agrformet.2017.10.029
18. C. Dow *et al*., Warm springs alter timing but not total growth of temperate deciduous trees. *Nature* **608**, 552–557 (2022). doi: 10.1038/s41586-022-05092-3; pmid: 35948636
19. C. Körner *et al*., Where, why and how? Explaining the low-temperature range limits of temperate tree species. *J. Ecol.* **104**, 1076–1088 (2016). doi: 10.1111/1365-2745.12574
20. Y. S. H. Fu *et al*., Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7355–7360 (2014). doi: 10.1073/pnas.1321727111; pmid: 24799708
21. C. M. Zohner, A. Rockinger, S. S. Renner, Increased autumn productivity permits temperate trees to compensate for spring frost damage. *New Phytol.* **221**, 789–795 (2019). doi: 10.1111/nph.15445; pmid: 30240028
22. Y. H. Fu *et al*., Larger temperature response of autumn leaf senescence than spring leaf-out phenology. *Glob. Change Biol.* **24**, 2159–2168 (2018). doi: 10.1111/gcb.14021; pmid: 29245174
23. C. M. Zohner, S. S. Renner, Ongoing seasonally uneven climate warming leads to earlier autumn growth cessation in deciduous trees. *Oecologia* **189**, 549–561 (2019). doi: 10.1007/s00442-019-04339-7; pmid: 30684009
24. I. Beil, J. Kreyling, C. Meyer, N. Lemcke, A. V. Malyshev, Late to bed, late to rise—Warmer autumn temperatures delay spring phenology by delaying dormancy. *Glob. Change Biol.* **27**, 5806–5817 (2021). doi: 10.1111/gcb.15858; pmid: 34431180
25. M. J. Paul, C. H. Foyer, Sink regulation of photosynthesis. *J. Exp. Bot.* **52**, 1383–1400 (2001). doi: 10.1093/jexbot/52.360.1383; pmid: 11457898
26. Y. Vitasse *et al*., Impact of microclimatic conditions and resource availability on spring and autumn phenology of temperate tree seedlings. *New Phytol.* **232**, 537–550 (2021). doi: 10.1111/nph.17606; pmid: 34235742
27. W. Buermann *et al*., Widespread seasonal compensation effects of spring warming on northern plant productivity. *Nature* **562**, 110–114 (2018). doi: 10.1038/s41586-018-0555-7; pmid: 30283105
28. C. Bigler, Y. Vitasse, Premature leaf discoloration of European deciduous trees is caused by drought and heat in late spring and cold spells in early fall. *Agric. For. Meteorol.* **307**, 108492 (2021). doi: 10.1016/j.agrformet.2021.108492
29. P. O. Lim, H. J. Kim, H. G. Nam, Leaf senescence. *Annu. Rev. Plant Biol.* **58**, 115–136 (2007). doi: 10.1146/annurev.arplant.57.032905.105316; pmid: 17177638
30. Y. Fracheboud *et al*., The control of autumn senescence in European aspen. *Plant Physiol.* **149**, 1982–1991 (2009). doi: 10.1104/pp.108.133249; pmid: 19201914
31. N. Jiang *et al*., Warming does not delay the start of autumnal leaf coloration but slows its progress rate. *Glob. Ecol. Biogeogr.* **31**, 2297–2313 (2022). doi: 10.1111/geb.13581
32. M. L. Goulden, J. W. Munger, S. M. Fan, B. C. Daube, S. C. Wofsy, Exchange of carbon dioxide by a deciduous forest: Response to interannual climate variability. *Science* **271**, 1576–1578 (1996). doi: 10.1126/science.271.5255.1576
33. C. M. Zohner, S. S. Renner, V. Sebald, T. W. Crowther, How changes in spring and autumn phenology translate into growth-experimental evidence of asymmetric effects. *J. Ecol.* **109**, 2717–2728 (2021). doi: 10.1111/1365-2745.13682
34. M. Friedl, J. Gray, D. Sulla-Menashe, MCD12Q2 MODIS/Terra +Aqua Land Cover Dynamics Yearly L3 Global 500m SIN Grid V006, NASA EOSDIS Land Processes DAAC (2019); https://doi.org/10.5067/MODIS/MCD12Q2.006.
35. B. Templ *et al*., Pan European Phenological database (PEP725): A single point of access for European data. *Int. J. Biometeorol.* **62**, 1109–1113 (2018). doi: 10.1007/s00484-018-1512-8; pmid: 29455297
36. J. O'Keefe, Phenology of Woody Species at Harvard Forest since 1990, v. 33, Environmental Data Initiative (2021); https://doi.org/10.6073/pasta/91e3b7c2548a0f2e251729eeacbce312.
37. G. Pastorello *et al*., The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci. Data* **7**, 225 (2020). doi: 10.1038/s41597-020-0534-3; pmid: 32647314
38. Materials and methods are available as supplementary materials.
39. S. Running, Q. Mu, M. Zhao, MOD17A2H MODIS/Terra Gross Primary Productivity 8-Day L4 Global 500m SIN Grid V006, NASA EOSDIS Land Processes DAAC (2015); https://doi.org/10.5067/MODIS/MOD17A2H.006.
40. Z. Wu *et al*., Atmospheric brightening counteracts warming-induced delays in autumn phenology of temperate trees in Europe. *Glob. Ecol. Biogeogr.* **30**, 2477–2487 (2021). doi: 10.1111/geb.13404
41. S. Fatichi, S. Leuzinger, C. Körner, Moving beyond photosynthesis: From carbon source to sink-driven vegetation modeling. *New Phytol.* **201**, 1086–1095 (2014). doi: 10.1111/nph.12614; pmid: 24261587
42. F. Pantin, T. Simonneau, B. Muller, Coming of leaf age: Control of growth by hydraulics and metabolics during leaf ontogeny.

*New Phytol.* **196**, 349–366 (2012). doi: 10.1111/j.1469-8137.2012.04273.x; pmid: 22924516
43. Y. Vitasse *et al.*, Elevational adaptation and plasticity in seedling phenology of temperate deciduous tree species. *Oecologia* **171**, 663–678 (2013). doi: 10.1007/s00442-012-2580-9; pmid: 23306445
44. A. J. Elmore, D. M. Nelson, J. M. Craine, Earlier springs are causing reduced nitrogen availability in North American eastern deciduous forests. *Nat. Plants* **2**, 16133 (2016). doi: 10.1038/nplants.2016.133; pmid: 27618399
45. R. E. Mason *et al.*, Evidence, causes, and consequences of declining nitrogen availability in terrestrial ecosystems. *Science* **376**, eabh3767 (2022). doi: 10.1126/science.abh3767; pmid: 35420945
46. P. M. Groffman *et al.*, Nitrogen oligotrophication in northern hardwood forests. *Biogeochemistry* **141**, 523–539 (2018). doi: 10.1007/s10533-018-0445-y
47. Q. Liu *et al.*, Delayed autumn phenology in the Northern Hemisphere is related to change in both climate and spring phenology. *Glob. Change Biol.* **22**, 3702–3711 (2016). doi: 10.1111/gcb.13311; pmid: 27061925
48. G. Liu, X. Chen, Y. Fu, N. Delpierre, Modelling leaf coloration dates over temperate China by considering effects of leafy season climate. *Ecol. Modell.* **394**, 34–43 (2019). doi: 10.1016/j.ecolmodel.2018.12.020
49. X. Lu, T. F. Keenan, No evidence for a negative effect of growing season photosynthesis on leaf senescence timing. *Glob. Change Biol.* **28**, 3083–3093 (2022). doi: 10.1111/gcb.16104; pmid: 35174579
50. L. Marqués *et al.*, Acclimation of phenology relieves leaf longevity constraints in deciduous forests. *Nat. Ecol. Evol.* **7**, 198–204 (2023). doi: 10.1038/s41559-022-01946-1; pmid: 36635342
51. C. Wu *et al.*, Contrasting responses of autumn-leaf senescence to daytime and night-time warming. *Nat. Clim. Chang.* **8**, 1092–1096 (2018). doi: 10.1038/s41558-018-0346-z
52. X. Zhang, M. Friedl, G. Henebry, VIIRS/NPP Land Cover Dynamics Yearly L3 Global 500m SIN Grid V001, NASA EOSDIS Land Processes DAAC (2020); https://doi.org/10.5067/VIIRS/VNP22Q2.001.
53. M. Rodell *et al.*, The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.* **85**, 381–394 (2004). doi: 10.1175/BAMS-85-3-381
54. B. Smith *et al.*, Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences* **11**, 2027–2054 (2014). doi: 10.5194/bg-11-2027-2014
55. C. Signarbieux *et al.*, Asymmetric effects of cooler and warmer winters on beech phenology last beyond spring. *Glob. Change Biol.* **23**, 4569–4580 (2017). doi: 10.1111/gcb.13740; pmid: 28464396
56. C. M. Zohner *et al.*, Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice, version 1.1, Zenodo (2023); https://doi.org/10.5281/zenodo.7797289.

## ACKNOWLEDGMENTS

We thank C. Körner and V. Sebald for comments on the manuscript. European ground data were provided by the members of the PEP725 project. **Funding:** C.M.Z. and L.Mi. were funded by the Ambizione grant PZ00P3_193646. T.W.C. was funded by DOB Ecology and the Bernina Foundation. D.P. was supported by the Faculty of Science Foundation of Charles University and the Hlávka Foundation. Y.H.F. was supported by the National Science Fund for Distinguished Young Scholars (42025101). B.D.S. and Y.V. were funded by the Swiss National Science Foundation grants PCEFP2_181115 and 315230_192712, respectively. **Author contributions:** C.M.Z. conceived of and developed the study, conducted the analyses, and wrote the manuscript. L.Mi. contributed to the development of the study and the remote-sensing analysis. R.B., D.P., and C.M.Z. conducted the experiments. L.Mo contributed to the analysis. D.R. contributed to the second experiment and the flux tower analysis. T.W.C., S.S.R., L.Mi., Y.V., Y.H.F., and B.D.S. contributed to the text. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All source code and data are available in Zenodo (*56*). The PEP725 data (*35*) used for this study are freely available at www.PEP725.eu. The satellite-derived phenology observations (*34*) are freely available at https://doi.org/10.5067/MODIS/MCD12Q2.006. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf5098
Materials and Methods
Figs. S1 to S27
References (*57*–*74*)
MDAR Reproducibility Checklist

Submitted 27 October 2022; resubmitted 3 February 2023
Accepted 23 May 2023
10.1126/science.adf5098

RESEARCH ARTICLE SUMMARY

MICROBIAL ECOLOGY

# Metabolic interaction models recapitulate leaf microbiota ecology

**Martin Schäfer†, Alan R. Pacheco†, Rahel Künzler, Miriam Bortfeld-Miller, Christopher M. Field, Evangelia Vayena, Vassily Hatzimanikatis, Julia A. Vorholt***

**INTRODUCTION:** Microbiome composition is tied to host health and ecosystem function. However, the processes that determine species abundances in an environmental context remain poorly understood. Given that microbiomes are bounded by resource availability, insights into the metabolic capacities of their constituent organisms present a key avenue for predicting interspecies interactions and broader microbiome assembly rules.

Plants constitute the largest biomass on Earth and are colonized by a variety of microorganisms that affect their health and growth. The composition of plant microbiomes has been found to be largely deterministic, suggesting the presence of defined drivers of community assembly. However, the metabolic mechanisms of the interactions underlying these patterns remain poorly understood.

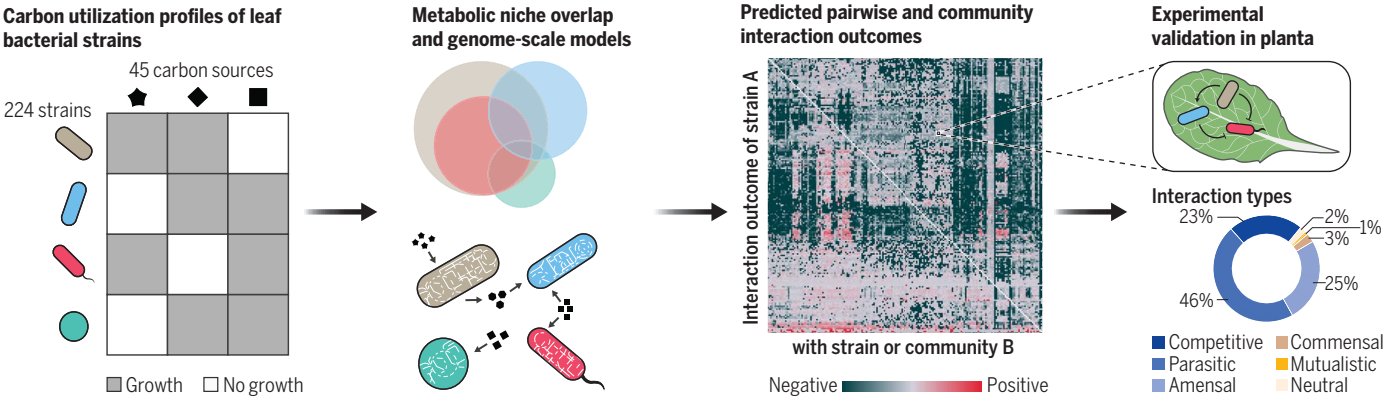
**RATIONALE:** To assess the degree to which patterns of resource allocation contribute to microbial interactions, we first profiled the carbon source–utilization capabilities of 224 representative bacterial strains isolated from leaves of wild *Arabidopsis thaliana* plants. To do this, we cultivated each isolate on 45 different carbon sources and developed a computational tool to score the growth of each strain in an automated way. This screen allowed us to calculate the degree of metabolic niche overlap between the strains, which informed predictions of resource competition on the plant host. We further leveraged this experimental data to generate a collection of 224 genome-scale metabolic models, which encompass the metabolic networks of each organism. We used these models to predict the outcomes of interspecies interactions on the plant host and to obtain mechanistic insight into specific patterns of resource use and exchange between these bacteria.

**RESULTS:** We found that the carbon source–utilization profiles of the strains exhibited a strong phylogenetic signature, on the basis of both the experimental screen and the pathways represented in the genome-scale models. We validated the performance of the models and of predictions based on niche overlap by conducting competition experiments with two sets of strains on the *Arabidopsis* host. These experiments revealed a dominance of negative interaction outcomes (i.e., a strain reached a lower overall population level upon cocolonization compared with monoassociation), which was in agreement with predictions of high interstrain niche overlap. The genome-scale models provided an additional degree of insight into these interactions, also corr[illegible] predicting instances of positive outc[illegible] observed in planta and further underscoring the importance of carbon metabolism to community assembly. After this experimental validation, we modeled interaction outcomes for more than 17,500 pairs of strains. We predicted that in 94% of pairings, at least one strain would experience a reduction in abundance compared with monoassociation. We then analyzed the metabolic fluxes underlying these predicted outcomes, which revealed a high prevalence of competition for sugars. This competition could be offset by increased uptake of amino and organic acids in our simulations, which points to the metabolic mechanisms underpinning the positive interaction outcomes that we observed in planta.

**CONCLUSION:** Our results indicate a major role for carbon source availability and metabolic interactions in community assembly in the plant host, which, together with the trait conservation observed here, likely contribute to the overall deterministic outcomes of plant microbiome assembly. In addition to recapitulating interspecies interactions in situ, our modeling framework represents a powerful tool for determining the metabolic mechanisms that lead to the emergence of specific ecological patterns. This knowledge will ultimately enable targeted microbiome design, which is pivotal to microbiome applications in health, agriculture, and the environment. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: jvorholt@ethz.ch
†These authors contributed equally to this work.
Cite this article as M. Schäfer *et al.*, *Science* **381**, eadf5121 (2023). DOI: 10.1126/science.adf5121

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adf5121




**Modeling of metabolic interactions predicts ecological outcomes for leaf microbiota members in situ.** We mapped the resource-use capabilities of 224 representative bacterial strains of the *A. thaliana* leaf microbiome by testing their ability to grow on 45 different carbon sources. These data allowed us to predict interaction outcomes between these strains using genome-scale models and metrics of niche overlap. Experimental validation in planta underscored the importance of resource competition and suggested metabolic mechanisms underlying specific interaction outcomes.

RESEARCH ARTICLE

MICROBIAL ECOLOGY

# Metabolic interaction models recapitulate leaf microbiota ecology

Martin Schäfer[1]†, Alan R. Pacheco[1]†, Rahel Künzler[1], Miriam Bortfeld-Miller[1], Christopher M. Field[1], Evangelia Vayena[2], Vassily Hatzimanikatis[2], Julia A. Vorholt[1]*

Resource allocation affects the structure of microbiomes, including those associated with living hosts. Understanding the degree to which this dependency determines interspecies interactions may advance efforts to control host-microbiome relationships. We combined synthetic community experiments with computational models to predict interaction outcomes between plant-associated bacteria. We mapped the metabolic capabilities of 224 leaf isolates from *Arabidopsis thaliana* by assessing the growth of each strain on 45 environmentally relevant carbon sources in vitro. We used these data to build curated genome-scale metabolic models for all strains, which we combined to simulate >17,500 interactions. The models recapitulated outcomes observed in planta with >89% accuracy, highlighting the role of carbon utilization and the contributions of niche partitioning and cross-feeding in the assembly of leaf microbiomes.

Various hosts, including animals (*1*, *2*), humans (*3*, *4*), and plants (*5*), support microbial communities with hundreds to thousands of different species. Recent studies have revealed that a deterministic relationship exists between environmental composition and community structure, even for complex microbiomes (*6*–*10*). Thus, a thorough understanding of the resources available to a community could enable prediction of its species and functional composition. Not only would the ability to make such predictions help us better understand the relationship between environment and phenotype, but it would also provide an accessible way to rationally design synthetic communities with defined functions.

Despite this potential, we still lack a comprehensive understanding of the information necessary to predict whether a specific organism will survive or be outcompeted in a particular environmental context. Resolving this question remains an active area of research, with studies that use interaction outcomes between pairs of organisms to predict overall community behavior representing a particularly attractive approach (*11*–*13*). Nonetheless, conducting the necessary mapping of possible interactions between all organisms remains experimentally challenging even for relatively simple communities and largely out of reach for complex host-associated microbiomes in situ.

Given these limitations, one may ask which ecological concepts can be used to predict the role of individual organisms within a community context. Microbial populations are fundamentally bounded by resource availability, where limiting amounts of micronutrients such as carbon, nitrogen, and phosphorus act to constrain community size and lead to the emergence of competition between organisms (*14*–*16*). Although the prevalence of these competitive interactions has begun to be quantified in a variety of ecosystems (*12*, *17*–*20*), the general rules by which they influence community-wide assembly patterns remain unresolved. As such, a commonly used approach is to infer the competitive potential between organisms from the degree of metabolic niche overlap between them, which relies solely on knowing their individual resource-utilization capabilities (*21*, *22*). Although a resource-by-resource understanding of the metabolic profiles of complex microbiomes can be challenging to obtain, it represents a more accessible approach for inferring interspecies interactions and has proven successful at predicting high-level patterns of species diversity in communities (*10*, *23*–*26*).

Despite its benefits, the predictive power of niche overlap is limited in that it cannot account for the emergence of positive ecological interactions between organisms either through resource partitioning, cross-feeding, orthogonal resource-acquisition strategies, or evolved cooperation (*27*–*30*). As a complementary approach, genome-scale metabolic models present a tractable way to integrate the effects of these additional mechanisms into predictions of community structure (*31*, *32*). Genome-scale models are mathematical representations of the metabolic capabilities of individual organisms. They incorporate genes, reactions, and metabolites associated with a given organism's metabolic network and are applied to quantitatively assess how organisms can use available resources for growth (*32*). In addition to generating predictions of resource allocation for individual organisms, combinations of genome-scale models of different organisms have been increasingly used to mechanistically describe pairwise and community-wide dynamics (*33*–*35*). Combining genome-scale models with predictions of niche overlap therefore presents a powerful way to predict interspecies interactions within microbiomes.

In this study, we generated metrics of metabolic niche overlap and a collection of genome-scale models to predict interaction outcomes between 224 bacterial members of the *Arabidopsis thaliana* leaf microbiome. This environment, known as the phyllosphere, is an oligotrophic habitat where competitive interactions between its resident microbes are prevalent (*19*, *36*, *37*). Because it contains a variety of different carbon sources (*38*, *39*), it is an ecosystem that is well suited for studying the effects of resource allocation on interspecies interactions and community structure. In particular, the exposed nature of the phyllosphere makes it a relatively accessible setting for study in controlled conditions. The microbiota of the phyllosphere has been shown to confer beneficial functions to the plant host (*40*–*42*), which, together with the scale of colonizable surfaces presented by leaves (*37*, *43*, *44*) and the economic importance of crops (*45*, *46*), underscores the value of studying the processes underlying its assembly (*47*, *48*).

## Results

### *Profiling the carbon source–utilization capabilities of phyllosphere bacteria*

Comprehensive collections of environmental strains are valuable resources for studying relevant host-microbe and microbe-microbe interactions (*42*, *49*–*52*). To examine interaction outcomes among members of the *Arabidopsis* phyllosphere (Fig. 1), we used the *At*-LSPHERE: a collection of 224 strains isolated from leaves of wild *A. thaliana* plants that represent a cross section of the taxonomic and functional diversity of the phyllosphere microbiota (*53*). We first assessed the ability of each strain to grow on solidified minimal medium supplemented with individual carbon sources from a set of 45 (table S1). These carbon sources, comprising a range of sugars, organic acids, sugar alcohols, one-carbon compounds, aromatic compounds, and amino acids, were selected on the basis of the high prevalence of resources such as glucose, sucrose, and some amino acids that are known to account for a substantial fraction of carbon available to leaf epiphytes (*36*, *38*, *39*, *54*). These carbon sources were also selected in part on the basis of the known ability of certain leaf strains to metabolize aromatic compounds, cleave glycosidic bonds,

[1]Institute of Microbiology, ETH Zurich, Zurich, Switzerland. [2]Laboratory of Computational Systems Biotechnology, École Polytechnique Fédérale de Lausanne, EPFL, Lausanne, Switzerland.
*Corresponding author. Email: jvorholt@ethz.ch
†These authors contributed equally to this work.

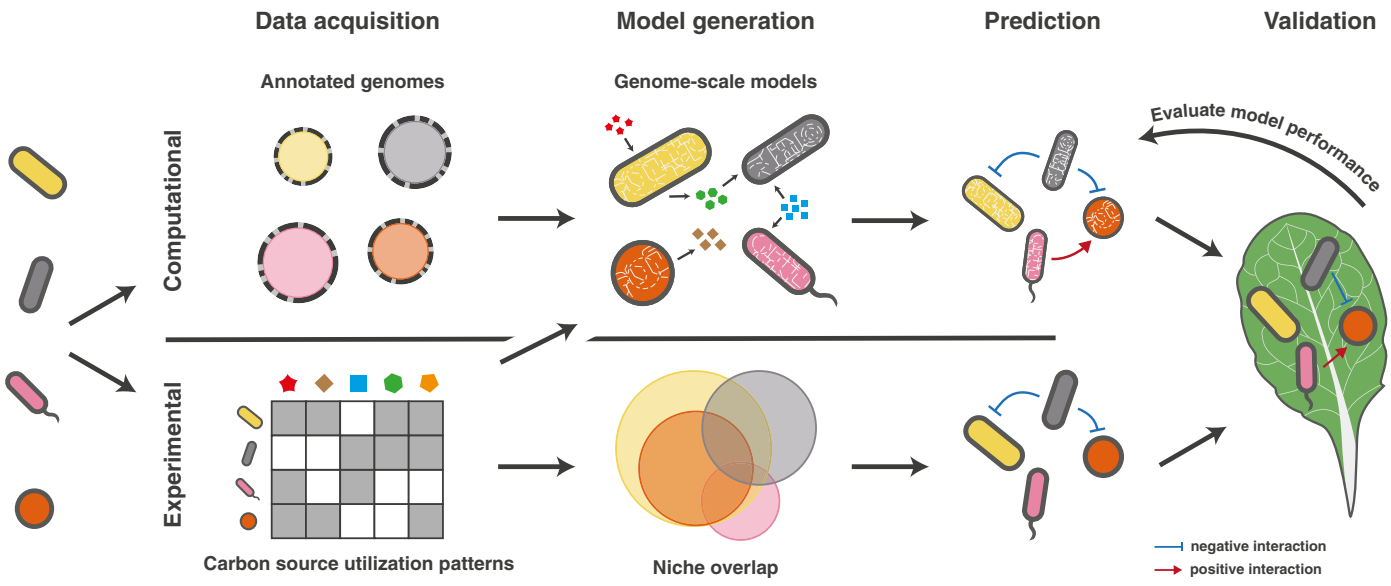



**Fig. 1. Overview of experimentally and computationally guided prediction and testing of metabolic interactions among bacterial members of the *Arabidopsis* phyllosphere microbiota.**

and/or utilize one-carbon compounds (*37*, *55*). In addition to these carbon sources, we supplemented the media with vitamins to enable the growth of auxotrophs (the set of carbon sources included amino acids, so we did not supply additional amino acids to the media). We used a plate assay to evaluate the growth of a given strain on a particular carbon source, which was compared with a carbon-free control, using both manual inspection and automated image processing (Materials and methods).

Our in vitro metabolic screen revealed phylogenetically contingent patterns of carbon-utilization capabilities across the *At*-LSPHERE collection (Fig. 2 and fig. S1). We found that strains belonging to the same genus displayed similar—or even identical—carbon source–utilization profiles, suggesting trait conservation between closely related species (Fig. 2A). We also identified genus-specific signatures of resource utilization, as well as cases in which strains could not grow on any of the tested carbon sources (e.g., *Chryseobacterium* and *Exiguobacterium*). These latter strains most likely have nutritional requirements that are not met by the minimal medium we provided (e.g., amino acid auxotrophies), because they were able to produce visible colonies on a complex medium (R-2A+M, Materials and methods). By contrast, strains belonging to the genera *Arthrobacter*, *Pseudomonas*, and *Rhizobium* grew on a large number of carbon sources [27.5 ± 6.9, 27.1 ± 3.7, and 25.8 ± 3.6, respectively (mean ± SD)]. To quantify these differences, we assigned a metric of substrate versatility, $V$, to each strain, defined as the percentage of all 44 growth-yielding carbon sources that each organism was able to use for growth (for this analysis, methylamine was excluded because it did not support the growth of any strain). The most versatile strain was *Arthrobacter* sp. Leaf145 ($V = 77\%$), followed by *Pseudomonas* spp. Leaf15 and Leaf98, and *Rhizobium* sp. Leaf202 ($V = 73$, 70, and 68%, respectively). When comparing the phylogenetic distribution of strain-specific substrate versatilities, we observed that *Beta-* and *Gammaproteobacteria* had above-average versatilities (Fig. 2B), whereas *Actinobacteria* (with the exception of *Arthrobacter*) and *Bacteroidetes* had below-average versatilities. Furthermore, our screen revealed the presence of a canonical niche occupied by methylotrophs that was distinct from and smaller than that of many other strains (Fig. 2A and fig. S2). Indeed, the lower average versatility of *Alpha-* and *Betaproteobacteria*—which include *Methylobacterium* and *Methylophilus* spp., respectively—relative to *Gammaproteobacteria* reflects the low versatility of methylotrophs.

Our screen also revealed the varying degrees to which each resource promotes growth (Fig. 2C). To quantify these patterns, we assigned a measure of substrate fertility, $F$, to each tested carbon source, defined as the percentage of strains that was able to grow on that particular carbon source. The most fertile or commonly used carbon source was glucose ($F = 81\%$ of strains), followed by succinic acid ($F = 80\%$) and glutamate ($F = 78\%$) (Fig. 2C and table S2). Among the most rarely used carbon sources were aromatic compounds, including tryptophan ($F = 9\%$) and coniferyl alcohol ($F = 8\%$), a common component of lignin.

Given the largely orthogonal metabolic niche occupied by methylotrophs relative to all other strains, we repeated our calculations of substrate fertility without methylotrophs to more clearly detect patterns of fertility among the other substrates. This analysis revealed that mono- and disaccharides build a core set of carbon sources commonly used by the large majority ($F \geq 75\%$) of tested strains and that the most fertile substrate, glucose, was consumed by 95% of the strains. The next most common group of carbon sources (≥ 50%) included organic acids and amino acids that feed into the tricarboxylic acid (TCA) cycle (fig. S3).

These experimentally determined carbon source–utilization capabilities allowed us to calculate the degree of niche overlap among our strains. We defined a niche overlap index (NOI) for each focal strain, A, with respect to any other strain, B, as the set of usable carbon sources shared by the focal strain and the competitor divided by the total number of carbon sources used by the focal strain (*23*, *56*):

$$\mathrm{NOI}_{A,B} = \frac{N_{\mathrm{C\ Sources,\ A} \cap \mathrm{B}}}{N_{\mathrm{C\ Sources,\ A}}}$$

We calculated NOI values for all pairs of strains, setting a threshold of 75% as likely prone to competitive exclusion (*57*) (Fig. 2D). In accordance with the similar carbon source–utilization profiles, we found high niche overlap within most genera, suggesting frequent intragenus competitive interactions. Additionally, we found that methylotrophic strains form two interconnected interaction spaces: one comprising *Methylobacterium* and the other comprising *Methylophilus* strains, underscoring

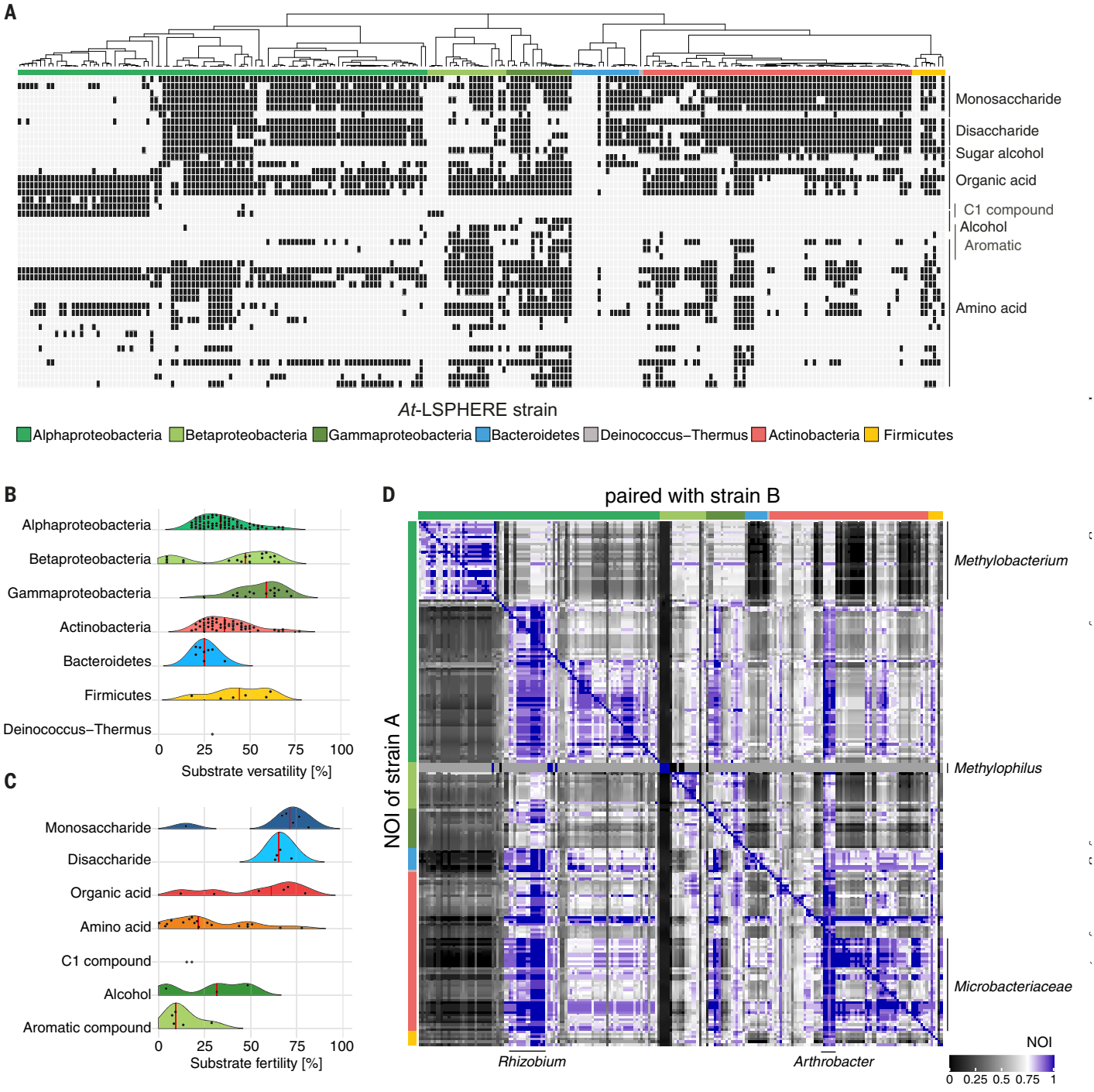



**Fig. 2. Carbon source–utilization screen reveals hotspots of high nutritional niche overlap.** (**A**) Carbon source–utilization map of 224 *At*-LSPHERE strains. Fill color of the boxes indicates growth (black) or no growth (gray) for each strain on the *x* axis on a given carbon source on the *y* axis. Top-level clustering for strains reflects phylogeny based on full-length 16*S* ribosomal RNA gene sequences. Colored bars indicate phylum or *Proteobacterium* class and correspond to the order in the legend. Carbon sources are sorted by compound groups. (**B** and **C**) Density plots showing (B) strain-specific substrate versatility grouped by phylum or *Proteobacterium* class (corresponding to the order in the legend) and (C) substrate fertility grouped by compound group. Black dots indicate individual values, and the red vertical line shows the median for each group. Substrates in classes C1 and aromatic compounds that also fall into another substrate group are only shown once and omitted from the other group. (**D**) Heatmap of niche overlap indices (NOI) for all binary combinations of 215 *At*-LSPHERE strains. The color of the tile indicates high (>0.75, blue shades) or low (<0.75, gray shades) degree of niche overlap for the strain on the *y* axis in combination with the strain on the *x* axis. Strains are sorted by phylogeny, and the colored bars indicate the phylum or *Proteobacterium* class. NOI was only calculated for strains that grew on ≥1 carbon source in vitro.

their degree of metabolic specialization compared with all other strains. Moreover, we found that members of the *Microbacteriaceae* had high niche overlap with a high number of other taxa, largely owing to their low substrate versatilities. We also observed the opposite case: *Rhizobium* strains had a low NOI with other taxa, whereas essentially all strains (except *Methylobacteria*) had a high NOI with *Rhizobium* strains.

### An in silico representation of phyllosphere bacterial communities

Having screened the carbon source–utilization capabilities of the *At*-LSPHERE collection, we sought to model how these bacteria would interact metabolically in a leaf environment. To do this, we adopted a stoichiometric metabolic modeling approach because it allowed us to explicitly simulate strain- and substrate-specific patterns of resource utilization, conversion, and exchange in a multitude of defined environmental conditions. We thus began by creating draft genome-scale metabolic reconstructions (*58*) for each member of the *At*-LSPHERE. This process yielded a separate metabolic network for each of the 224 *At*-LSPHERE strains, containing all the metabolic reactions predicted to be contained by each organism on the basis of its genome annotation. After generating these draft metabolic reconstructions, we used flux balance analysis (*31*) to simulate the growth of each strain individually within a complete medium containing all 45 carbon sources used in our in vitro screen. We found that whereas some of our reconstructions were predicted to produce biomass when using this complete medium, the majority could not do so when only individual carbon sources were provided (Fig. 3A). Although this inability to grow was at odds with our experimental results, it was not unexpected because models generated solely from genomic information are often missing key metabolic pathways and thus have limited predictive power without additional experimental curation (*33*, *59*–*61*). We therefore used the results of our carbon source screen to fill the apparent gaps in each metabolic reconstruction by using the tool NICEgame (*62*), which suggests biologically relevant alternative pathways to account for missing reactions by leveraging information on reaction thermodynamics. We followed this step with additional manual curation (Materials and methods, table S3, and fig. S4) and then compared each of the resulting 224 models with their corresponding strains on the basis of their ability (or inability) to use the 45 carbon sources selected for our in vitro screen (fig. S5). This process resulted in each genome-scale model having a final balanced accuracy of 0.98 ± 0.07 (Fig. 3A) (mean precision 1, mean recall 0.96), with 214 models (96%) having a balanced accuracy of at least 0.85, and 189 models (84%) being 100% accurate.

Our curation process yielded a collection of 224 genome-scale models that revealed further metabolic capabilities and physiological characteristics of the *At*-LSPHERE collection (Fig. 3). The models contained between 1568 (*Flavobacterium* sp. Leaf359) and 3004 (*Pseudomonas* sp. Leaf15) reactions (Fig. 3B), which moderately correlated with the genome size of their corresponding strain [coefficient of determination ($R^2$) = 0.54, fig. S6A]. Reactions involved in the biosynthesis of glycerophospholipids, key components of the cell membranes of Gram-negative bacteria, were enriched in members of the *Proteobacteria* within our collection (Fig. 3C). We also found *Firmicutes* and *Actinobacteria* to contain more reactions related to degradation of plant polysaccharides (e.g., melibiose, raffinose, and sucrose) relative to other taxonomic classes and identified an enrichment for benzoate degradation capabilities as previously reported in *Betaproteobacteria* (*63*, *64*). The high degree of similarity to our strains' known carbon source–utilization capabilities was supported by a clustering analysis of the reactions contained in each model (Fig. 3D). In this analysis, *Methylobacteriaceae* spp. and *Methylophilaceae* spp. formed clusters distinct from most other models, and *Acidovorax* strains clustered together with the majority of *Pseudomonas* and *Rhizobium* organisms. These clusters largely mirrored those generated from our in vitro data alone (Fig. 3E), which underscored the phylogenetic contingency of resource-utilization patterns and may underlie the deterministic assembly of microbiomes in general. With this in silico representation of the *At*-LSPHERE collection, we proceeded to use our genome-scale models to simulate the ecological outcomes of mixed-culture experiments.

### Computational models recapitulate ecological patterns in planta

We assessed the predictive power of the genome-scale models and the NOI against experimental data in two in planta experiments (Fig. 4, A and B). In a first experiment, we inoculated *A. thaliana* seedlings with pairwise combinations of seven bacterial strains, or with a community consisting of all seven strains (referred to as SynCom7), and compared the colonization level of each strain in combination with the colonization density achieved in monoassociation (Fig. 4B). To represent a range of metabolic capabilities, we selected strains across a gradient of substrate versatility values and hence a gradient of NOI. For the first experiment, we selected representative strains from the highly versatile *Arthrobacter* spp., *Pseudomonas* spp., and *Rhizobium* spp. (Leaf145, $V$ = 77%; Leaf15, $V$ = 73%; and Leaf202, $V$ = 68%; respectively), along with four additional strains with intermediate-to-low versatility values: *Rhodococcus* sp. Leaf233, *Sphingomonas* spp. Leaf34 and Leaf257, and *Microbacterium* sp. Leaf179 ($V$ = 55, 36, 45, and 41%, respectively). On the basis of NOI, we expected the strains with lower versatilities to more frequently decrease in abundance upon cocolonization with other strains, either in pairwise combinations or in a community context (Fig. 4C). However, all other strain pairings had low NOI values (<75%), limiting the degree to which we could infer interaction outcomes for these combinations.

We used the genome-scale models corresponding to these seven strains to simulate their growth within an environment reflecting the carbon source availability of the *Arabidopsis* phyllosphere (Materials and methods). To obtain an indicator of the ecological outcomes of these pairings, we compared the predicted growth of each strain on its own with that in coassociation with another: A strain was deemed to have experienced a negative interaction outcome if its biomass production rate in coassociation was lower than that on its own, and a positive outcome if its biomass production rate was higher in coassociation. These simulations predicted an approximately even distribution of positive and negative outcomes for all seven organisms (Fig. 4C). However, when we considered the magnitude of the changes in growth experienced by each strain, we found that predictions skewed more strongly negative than positive. This observation is consistent with previous descriptions of the prevalence of competitive interactions in the phyllosphere (*18*, *19*), which suggests the commonality of exclusion of one strain by another. Indeed, when we compared the growth of each strain on its own with that in the presence of all six other strains, we predicted almost exclusively negative outcomes. Despite this overall distribution, we found that the number of positive or negative outcomes differed depending on the strain in question, with *Microbacterium* sp. Leaf179 and *Sphingomonas* sp. Leaf257 almost always experiencing a reduction in growth when paired with another strain. By contrast, *Rhizobium* sp. Leaf202, *Pseudomonas* sp. Leaf15, and *Arthrobacter* sp. Leaf145 were most often predicted to benefit from coinoculation with another of our selected strains. These predicted positive outcomes corresponded to instances of low NOI values (Fig. 4C), suggesting that these strains could metabolically evade competition by a partner strain and may additionally benefit from changes in the ecosystem (e.g., cross-feeding) induced by the partner strain.

The interaction outcomes of these seven strains in planta were largely congruent with those predicted by NOI and metabolic modeling. Of the outcomes that showed a fold change

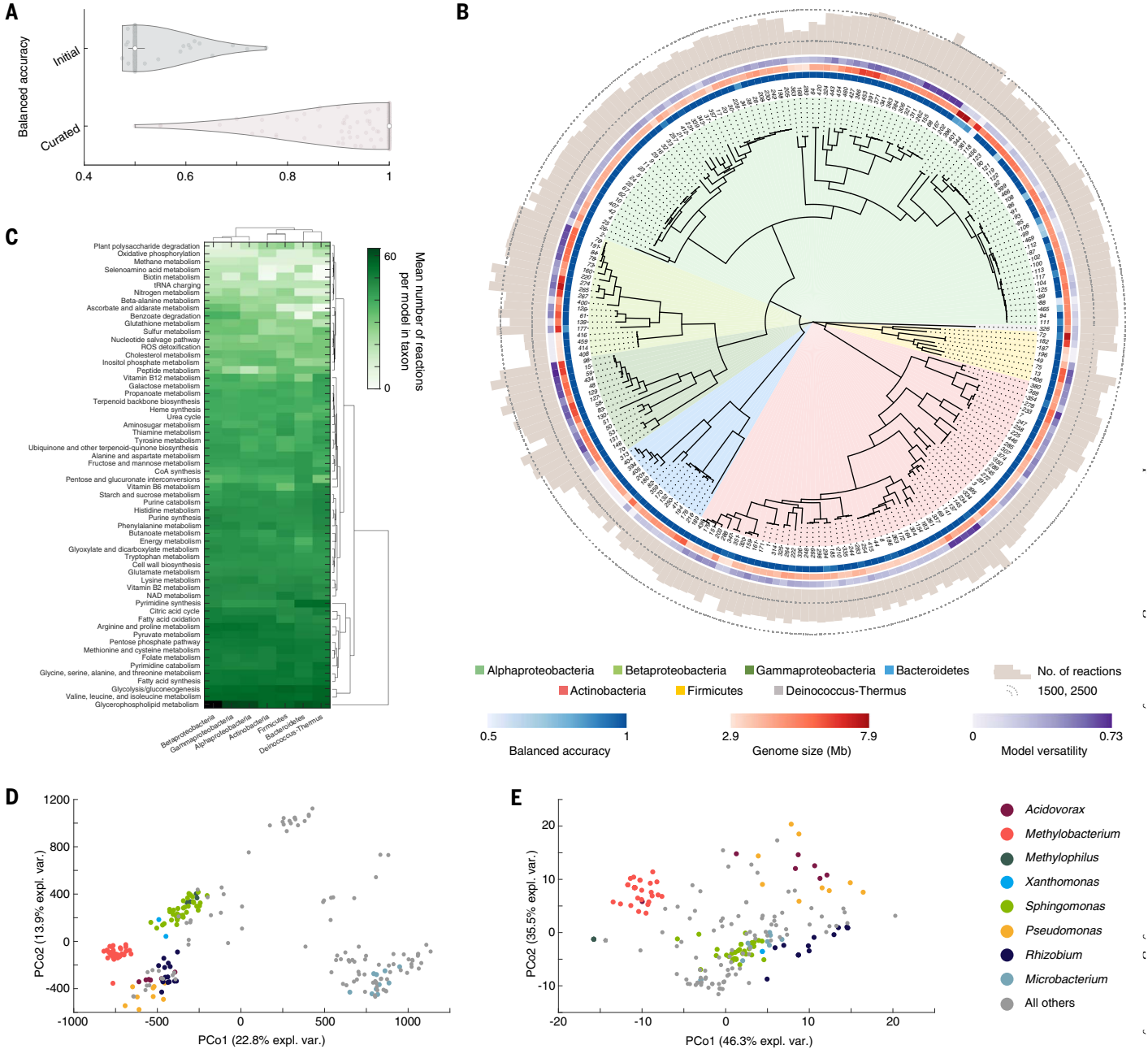



**Fig. 3. Overview of collection of metabolic models for 224 bacterial members of the *A. thaliana* leaf microbiota.** (**A**) Distributions of balanced accuracy as tested on 45 carbon sources for models generated solely from genomic information (top) and after curation (bottom). (**B**) Attributes of models as clustered by taxonomic phylum and class. Numbers bordering phylogenetic tree denote strain identity (e.g., 220 corresponds to *Xylophilus* sp. Leaf220). The innermost ring represents the balanced accuracy of each model, the second ring indicates the size of the genome used to generate each model (*53*), and the third ring represents the versatility, *V*, of each model. The outermost bars represent the number of reactions contained in each model (minimum = 1568, maximum = 3004). (**C**) Clustered heatmap of 60 most highly represented reaction types, scaled by the number of models corresponding to each phylum and class. (**D** and **E**) Principal coordinates analysis (PCo) of strains as determined by (D) reactions in genome-scale models and (E) in vitro carbon source screen, with select genera highlighted.

greater than 2 or less than 0.5, we found that a majority were negative, confirming the dominance of competition between these strains (Fig. 4D). In the SynCom7 condition, a significant abundance reduction was observed for all strains except for *Rhizobium* sp. Leaf202, matching our predictions and supporting the presence of increased competitive pressure in a community context. Our experimental data also highlighted the predictive power of NOI data alone for predicting interaction outcome directionality, with 93% of instances of negative outcomes corresponding to an NOI >0.75. Predictions generated by the genome-scale models yielded an increased degree of granularity, recapitulating the significant interaction directionalities that we observed in planta with a balanced accuracy of 89% (tables S4 and S5). These predictions additionally captured the weaker directionalities of border cases that did not meet our experimental

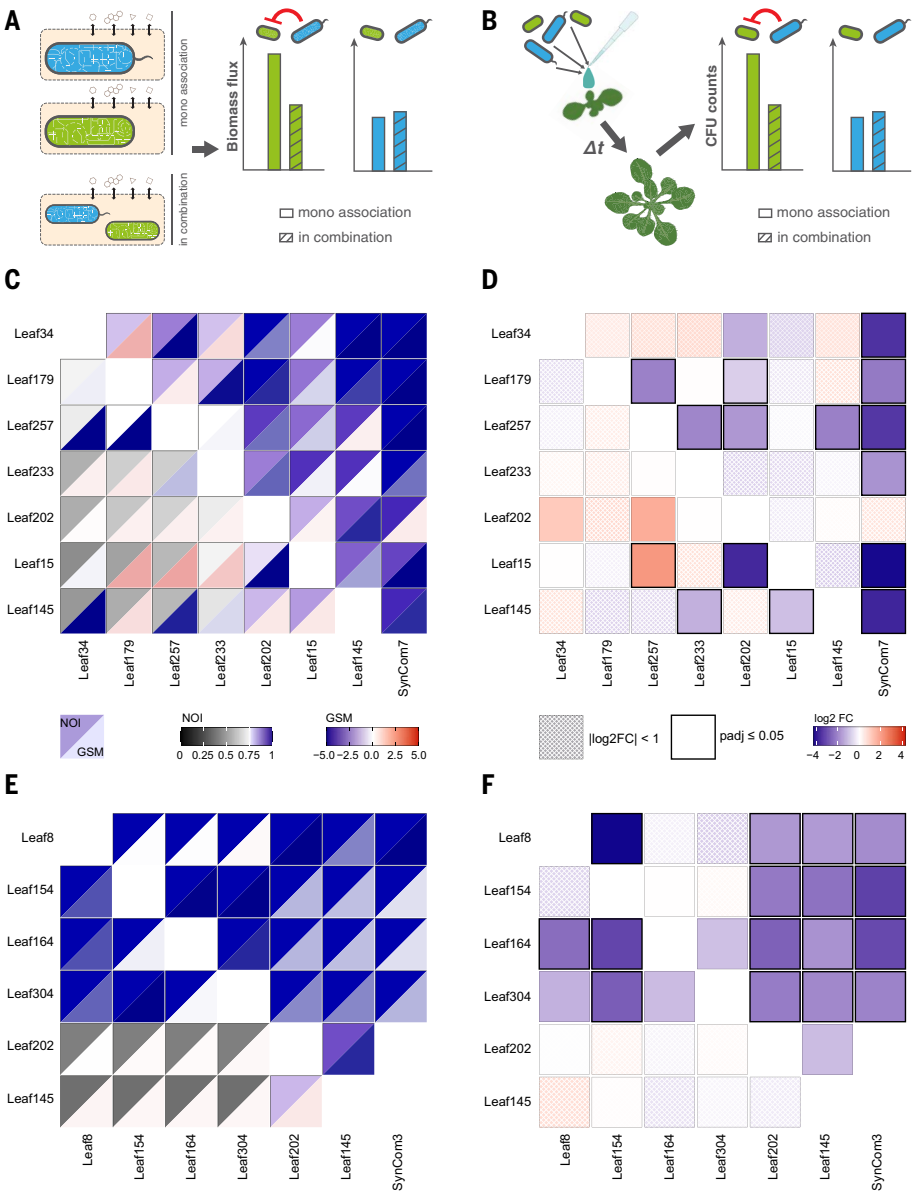



**Fig. 4. Model predictions and in planta validation of interactions.** (**A** and **B**) Schematic overview of (A) interaction simulations using genome-scale models and (B) phyllosphere inoculation and interaction mapping procedure. Interactions are inferred by comparing a strain's growth in combination with another strain with that in monoculture. This is computed as the $\log_2$ ratio of biomass fluxes for the genome-scale models or as the $\log_2$-fold change in CFUs per gram plant fresh weight for the in planta experiments. (**C** and **E**) Predicted interaction outcome for the strain indicated on the *y* axis in combination with the strain or SynCom indicated on the *x* axis, based on niche overlap index (NOI, top left) and genome scale models (GSM, bottom right). The fill color indicates high (>0.75, blue shades) or low (<0.75, gray shades) degree of niche overlap and predicted positive (red shades) or negative (blue shades) interactions based on genome-scale metabolic models. (**D** and **F**) Interaction outcomes observed in planta. Heatmap showing the $\log_2$-fold changes ($\log_2$FCs) (pairwise or SynCom inoculation versus monoassociation) for the strain on the *y* axis in combination with the strain or SynCom indicated on the *x* axis, based on absolute abundances obtained by CFU enumeration ($n = 11$ to 12). The color of the boxes reflects the observed $\log_2$FC, and the black frames around the boxes indicate a significant difference compared with the monoassociation condition (two-sided Wilcoxon rank sum test, Holm-adjusted $P \leq 0.05$). $|\text{Log}_2\text{FC}| < 1$ are overlaid with a crosshatched pattern. Confusion matrices comparing modeling predictions and in planta outcomes are provided in table S4, and colony counts for the in planta experiments are provided in table S5.

cutoff, as seen in the positive interaction outcome for *Sphingomonas* sp. Leaf34 paired with *Microbacterium* sp. Leaf179, of *Arthrobacter* sp. Leaf145 paired with *Rhizobium* sp. Leaf202, and of *Rhizobium* sp. Leaf202 in the SynCom7 condition. These weakly positive effects occurred despite high predicted NOI, suggesting the influence of additional metabolic mechanisms not encompassed by competition.

We sought to further validate our predictions with a second experiment spanning more extreme degrees of niche overlap. We generated pairings including the two highly versatile strains *Rhizobium* sp. Leaf202 and *Arthrobacter* sp. Leaf145 ($V = 68\%$ and 77%, respectively) as well as the low-versatility strains *Frigoribacterium* sp. Leaf8, *Curtobacterium* sp. Leaf154, *Rathayibacter* sp. Leaf164, and *Frondihabitans* sp. Leaf304 ($V = 27\%$ for all four strains). All pairings resulted in a high degree of niche overlap for the low-versatility strains, whereas the NOI for either Leaf202 or Leaf145 with any other strain was low. Correspondingly, the genome-scale models predicted almost exclusively negative interaction outcomes for all low-versatility strains, with weakly positive outcomes for Leaf202 in combination with all strains except Leaf145, as well as stronger positive outcomes for Leaf145 in combination with most other strains (Fig. 4E). We additionally predicted interaction outcomes for each of the three low-versatility strains paired with a combination of Leaf202 and Leaf145 (referred to as SynCom3), which resulted in strongly negative interaction outcomes for each strain. These predictions were tested with in planta experiments for these pairings: As predicted, all significant instances of $\log_2$-fold changes greater than 1 were negative, including those for each low-versatility strain in the SynCom3 condition (Fig. 4F). Additionally, Leaf202 and Leaf145 each experienced two instances of weakly positive outcomes when paired with other low-versatility strains, which were captured by our genome-scale modeling predictions. Our experiments confirmed the validity of the computational predictions (tables S4 and S5) and highlighted the strong contribution of resource competition to strain-specific interaction outcomes in situ.

To test the specificity of our predictions to a leaf environment, we carried out a series of in vitro cultivations in which we cultured the same strain pairs and communities in shake flasks (Materials and methods). These cultivations resulted in exclusively negative interaction outcomes for all strains in the SynCom7 and SynCom3 conditions (fig. S7), similarly to our in planta experiments. These outcomes, which are consistent with an additional set of community simulations that used randomly selected strains (fig. S8), highlight the increased degrees of competition that organisms may be subject to in multispecies settings. However,

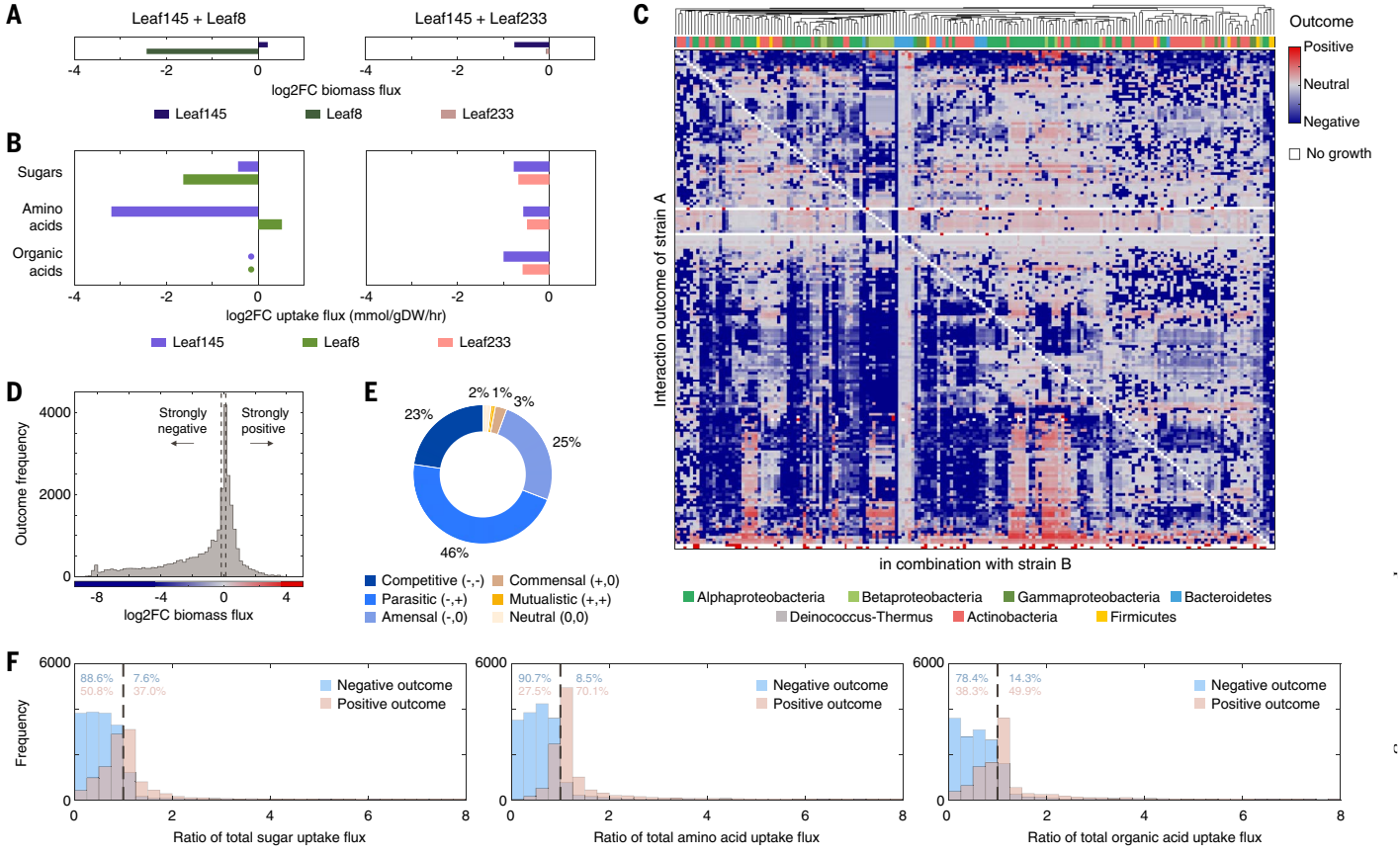



**Fig. 5. Model-predicted interaction outcomes and mechanisms.** (**A** and **B**) $Log_2$FCs of (A) biomass and (B) resource uptake fluxes for two representative interactions validated in planta. Dots indicate absolute $log_2$FCs of less than 0.05 mmol $gDW^{-1}$ $hour^{-1}$. (**C**) Predicted pairwise interaction outcomes between all 188 nonmethylotrophic strains in the *At*-LSPHERE ($n$ = 35,156 outcomes for 17,578 pairs). Hierarchical clustering was performed on interaction outcomes, with strain-specific phylogeny highlighted. White cells denote instances of no predicted growth in both mono- and coculture. (**D**) Distribution of pairwise interaction outcomes ($n$ = 35,156). Dashed lines separate outcomes in which a strain's predicted biomass flux in coculture was either less than 90% of that in monoculture (strongly negative), within 10% of that in monoculture (neutral), or more than 110% of that in monoculture (strongly positive). (**E**) Classification of pairwise ecological outcomes ($n$ = 17,578). (**F**) Distributions of flux ratios between resource uptake in coculture and monoculture, according to corresponding interaction outcome. Only simulations in which a strain achieved growth in both monoculture and coculture are considered ($n$ = 28,316 outcomes). Differences between uptake rates of resource types provided in the simulated medium are highlighted for sugars (left), amino acids (center), and organic acids (right). Distributions of uptake fluxes are statistically significant for all three resource types ($p \ll 1 \times 10^{-10}$) as determined by one-tailed Mann-Whitney $U$ tests. For clarity, horizontal axes are truncated and show 98.8% of outcomes for sugars, 98.7% for amino acids, and 94.7% for organic acids. The dashed line at ratio of 1 separates instances of lower or higher uptake flux between coculture and monoculture, with percentages highlighting the number of instances less than or greater than 1.

our in vitro experiments also resulted in substantially fewer positive outcomes of pairwise interactions when compared with our in planta results, suggesting competitive pressures inherent to batch cultures that are not captured by our use of NOI and genome-scale models. In particular, the lack of spatial structure may favor strains that experience fast growth when substrate availability is high and thus outcompete slower-growing strains within the timescale of a batch experiment. Moreover, this rapid depletion of nutrients contrasts with the resource dynamics of leaf surfaces, which exhibit a steady resupply of resources that can be accessed by epiphytic microbes (*38*). Although we do not explicitly consider spatial structure in our use of NOI and genome-scale models, these tools generate predictions based on a broader consideration of the various resources that can be used by the organisms at steady state. This assumption may therefore better reflect a broader and more continuously supplied pool of resources, which can be used by microbes on a population level in spatially structured settings.

### *Compensatory metabolic mechanisms offset resource competition*

In addition to predicting interaction outcome directionalities, we used genome-scale modeling to explore the metabolic mechanisms that could be underlying the observed ecological patterns. We first examined experimentally validated interactions to determine changes in resource uptake rates that emerged as a result of pairing two strains together. As representative examples, we looked specifically at two interactions involving *Arthrobacter* sp. Leaf145, a highly versatile strain that experienced a weakly positive effect when paired with *Frigoribacterium* sp. Leaf8 and a negative effect when paired with *Rhodococcus* sp. Leaf233 (Fig. 4, C to F, and Fig. 5A). Our flux balance simulations had predicted that in both cocultures, Leaf145 would have a lower net uptake flux of sugars compared with those experienced in monoculture (Fig. 5B). This was also the case for Leaf8, which was additionally predicted to experience a reduction in amino acid uptake flux when paired with Leaf145. Whereas a similar reduction in amino and organic acid uptake occurred for Leaf233, Leaf145 was able to take up amino acids at rates similar to those noted in monoculture. The contribution of this reallocation of resources to the dominance of Leaf145 can be seen in its interaction with Leaf8, in which Leaf145 was able to shift its

metabolism to take up a slightly greater quantity of amino acids, as occurred in monoculture. The increased availability of these resources, in part a product of the low metabolic activity of Leaf8, suggests metabolic cross-feeding as a contributing factor in the positive effect experienced by Leaf145 in coculture.

The emergence of distinct resource-allocation patterns between these strain pairs prompted us to ask how widespread they could be across a wider ecosystem. We thus used the genome-scale models to generate predictions of pairwise interaction outcomes for all nonmethylotrophic strains in the *At*-LSPHERE, carrying out a total of 17,578 pairwise simulations (Fig. 5C). These simulations revealed the high prevalence of competitive effects, with 63.2% of outcomes predicted to be negative (Fig. 5D). Of these, the large majority (representing 58.2% of all outcomes) were predicted to be strongly negative, which we defined as a strain's biomass production rate in coculture being less than 90% of that in monoculture. Conversely, 25.5% of all outcomes were predicted to be strongly positive (biomass production in coculture more than 110% of that in monoculture), further underscoring the strength of competitive pressures within leaf bacterial communities. We combined these strain-specific effects to define interaction outcomes between pairs, which revealed that in 94% of interactions, at least one participant experienced a negative outcome (Fig. 5E). As such, cooperative interactions were predicted to be relatively rare, with 3% of all pairwise interactions resulting in commensalism and only 1% resulting in mutualism.

An analysis of metabolic fluxes across our simulations revealed key differences in rates of resource uptake between positive and negative interaction outcomes (Fig. 5F). As expected, the models predicted significantly lower resource uptake rates for organisms that decreased in abundance relative to monoculture than for organisms that increased in abundance. However, whereas nearly all flux ratios associated with negative outcomes were below 1 (occurring for sugars, amino acids, and organic acids in 88.6, 90.7, and 78.4% of negative outcomes, respectively), many positive interaction outcomes displayed increases in amino acid and organic acid uptake flux (in 70.1 and 49.9% of positive outcomes, respectively). These patterns thus suggest a stronger degree of competition for sugars among phyllosphere bacteria, which is more likely to be offset by increased uptake of amino and organic acids.

These predicted compensatory mechanisms suggested that strains with higher substrate versatilities would be more likely to increase in abundance when in the presence of another strain. Indeed, strains with very high versatilities (e.g., *Acidovorax* spp. and *Pseudomonas* spp.) were predicted to have higher than average rates of experiencing positive outcomes (fig. S9). However, strain-specific substrate versatility itself was a poor predictor of positive outcome frequency ($R^2 = 0.10$) because strains with very low versatilities (e.g., *Chryseobacterium* spp., which did not grow on any of the supplied carbon sources individually, or *Bosea* sp. Leaf344) were also predicted to experience strongly positive outcomes in rare cases. These effects suggest a role of cross-fed amino or organic acids as valuable sources of metabolic complementation (*65*), which in addition to enabling highly versatile strains to outcompete a partner strain, can also facilitate the survival of specialized organisms.

## Discussion

Understanding the factors that contribute to the assembly of microbiomes remains a challenge for the study of natural ecosystems. In this study, we have shown how predictions based on the resource-utilization capabilities of individual bacteria enable recapitulation of interaction outcomes observed in the phyllosphere. These predictions underscored the prevalence of competitive interactions between leaf strains. Additionally, the high accuracy (Fig. 4) of these predictions suggested that competition for resources is responsible for many outcomes observed in situ, as also recently observed in community contexts (*19*). In particular, our experiments showed how high degrees of niche overlap were reliably predictive of negative interaction outcomes in planta. This relationship demonstrates the role that highly versatile organisms play in communities, as well as the implications for the design of microbial consortia based on cooperative interactions, where pairings of highly competitive organisms can be avoided with greater certainty. Our work highlights how a resource-by-resource understanding of an organism's catabolic potential in monoculture is sufficient for making high-level predictions of competition in an ecologically relevant setting.

In addition to mapping the metabolic capabilities of a large strain collection at high resolution, our in vitro carbon source screen informed the curation of genome-scale metabolic models for each of the 224 strains. Our experimentally curated collection of genome-scale models allowed us to go beyond the potential of niche overlap by predicting positive interaction outcomes and suggesting their molecular mechanisms. Our curation process provided a quantitative basis for the need to incorporate experimental data into model generation, given the low accuracy of the initial draft models, which were based on genome annotations alone. Although improved annotation resources, as well as organism-specific biomass compositions, may serve to reduce this source of uncertainty (*60*, *61*), we expect experimental curation to remain essential for selecting the optimal combination of gap-filled reactions to recapitulate an organism's exact resource-utilization profile.

Our metabolic modeling predicted interaction outcomes on the basis of carbon source–utilization capabilities. Further curation that integrates additional properties such as vitamin auxotrophies and storage (*66*), as well as metabolic shifts occurring from gene regulation (*67*, *68*), is likely to improve the quantitative predictive power of this framework (*33*, *69*). Moreover, although our predictions based on metabolic mechanisms are informative of key aspects underlying the assembly of plant-associated microbiomes, they do not consider additional factors that are known to shape leaf communities, such as signaling molecules, antagonistic interactions, or host immunity (*5*, *70*–*73*). A potential consequence of this limitation may be seen in the example of *Pseudomonas* sp. Leaf15, which experienced a strong negative interaction in planta when paired with *Rhizobium* sp. Leaf202 (Fig. 4, C and D). Even though this outcome was qualitatively predicted by our simulations, the presence of a type VI secretion system in Leaf202, previously described to be active against a *Pseudomonas* strain (*42*), may also be producing an additional inhibitory effect not captured in our modeling framework. Furthermore, although our experimental approach that quantified community composition on the basis of bulk sampling aligns with our use of flux balance analysis, it abstracts away the influence of spatial structure on interaction outcomes. Leaf surfaces are known to exhibit considerable nutrient heterogeneity, which, together with the aggregation of microbes around stomata or in microdroplets, is likely to affect interaction directions at the microscale (*36*, *74*, *75*). Capturing varying degrees of uncertainty in our modeling predictions can partially represent this heterogeneity (fig. S10), but an improved understanding of the nutrient compositions at specific locations on the leaf may enable future parametrization of existing metabolic modeling tools that can explicitly consider spatial structure (*76*–*79*).

Our present results underscore the strengths of an integrated approach for generating ecological predictions in a mechanistic yet scalable way, while establishing a computational resource for further exploration of the molecular mechanisms that underlie the assembly of complex microbiomes.

## Materials and methods

### *Cultivation of bacteria*

*At*-LSPHERE strains were grown on R-2A agar (Sigma-Aldrich) supplemented with 0.5% (v/v) methanol (R-2A+M) at 22°C. Strains were routinely streaked out from a cryo stock stored at −80°C, grown for 4 days, streaked on fresh

R-2A+M plates, and grown for 3 more days before the start of experiments. If necessary, streptomycin (20 μg $ml^{-1}$) was added to the medium for selection.

### Carbon source screen

The minimal medium agar plates were prepared with a common minimal medium composition (*80*). One liter of the final medium contained 2.4 g $K_2HPO_4$, 1.08 g $NaH_2PO_4 \cdot 2H_2O$, 1.62 g $NH_4Cl$ and 0.2 g $MgSO_4 \cdot 7H_2O$, and 15 g noble agar (Becton, Dickinson and Company). The medium was supplemented with the following trace elements: 15 mg $Na_2EDTA \cdot 2H_2O$, 3 mg $FeSO_4 \cdot 7H_2O$, 4.5 mg $ZnSO_4 \cdot 7H_2O$, 3 mg $CoCl_2 \cdot 6H_2O$, 0.64 mg $MnCl_2$, 1 mg $H_3BO_3$, 0.4 mg $Na_2MoO_4 \cdot 2H_2O$, 0.3 mg $CuSO_4 \cdot 5H_2O$, and 3 mg $CaCl_2 \cdot 2H_2O$, and vitamins: 500 μg D-pantothenic acid hemi calcium salt, 100 μg biotin, 400 μg riboflavin, 400 μg thiamine HCl, 200 μg pyridoxal HCl, 150 μg p-amino benzoic acid, 200 μg cobalamin, 50 μg lipoic acid, 150 μg nicotinic acid, and 100 μg folic acid. All media components were prepared with Milli-Q quality water (Millipore). Each carbon source was added as a 10× stock solution to the premixed medium, except for coniferyl alcohol and tyrosine, which were added to the medium directly as powder. The concentration of each carbon source was normalized to 30 mM carbon (e.g., 5 mM glucose and 30 mM methanol). A list of all 45 carbon sources and final concentrations can be found in table S1.

*At*-LSPHERE isolates grown on R-2A+M medium were suspended (1 μl loopful) in 1 ml 10 mM magnesium chloride solution, which corresponds to an optical density at 600 nm ($OD_{600}$) of approximately 0.1 to 0.3. The cell suspensions were transferred to 96-well plates before spotting on agar plates, leaving every other well empty in a checkerboard manner to reduce the risk of contamination and allow more space on the agar plate for each strain. Increasing the distance between spots also strongly reduces the risk of strain interactions based on secondary metabolite production (e.g., antibiotics) that could lead to false-negative results (*42*). Cell suspensions (2 μl) were spotted on top of each minimal medium plate containing individual carbon sources, as well as on a minimal medium plate without carbon source and a R-2A+M plate. Spotting was carried out using a Rainin Liquidator Manual Pipetting System (Mettler Toledo). Plates were dried under laminar flow. Bacteria were incubated at 22°C and photographs were taken at 5, 7, and 10 days after spotting. In sum, the screen comprised three total rounds including 96 strains each, and the last round also comprised 56 randomly selected strains that were screened a second time for validation purposes (table S7). Plate images are available through Zenodo (*81*).

### Analysis of bacterial growth in carbon source screen

Colony growth was scored for all strains after 7 days (0 = no growth, 1 = growth) by comparing growth on each carbon source with the control plate without a carbon source. We used a dual approach incorporating automated growth scoring based on pixel intensities, using a common threshold for all strains and manual scoring through visual inspection.

For the automated analysis, we developed the software tool platescan (github.com/MicrobiologyETHZ/platescan), which enables unbiased scoring of all strains over different conditions with a firm growth cutoff. Briefly, platescan uses cross-correlation techniques to crop the plate image, locate the colony grid layout, and then determine the best-fitting colony size and location on the basis of the assumption that they are approximately circular. The pixel intensities are rescaled between 1 and 99% of the intensity distribution to ensure that the images have approximately the same contrast. The program then reports foreground and background pixel intensities in red, green, blue, and grayscale, which are then used to threshold growth versus no growth. An example can be found on the program's GitHub repository.

We used platescan to assign growth values to each strain with the following parameters: -r:20, -p: 10,–min_r: 15,–max_r: 40,–edge: 200 40 200 40. In addition, two parameters were set individually for each screening round because of variation in picture zoom level. Screen1: -x 105 -y 108, screen2: -x 110 -y 109, screen3: -x 114 -y 113. We used the pixel intensities reported in the red channel with a defined minimum threshold to consider a strain as having grown. Pixel intensities were obtained by subtracting background values from foreground values for each colony and subsequently subtracting the no-carbon control value. For the few cases in which strains spread into the area used for background intensity calculation, the background value of the neighboring colony was subtracted.

For a fraction of strains, we also observed substantial growth without the addition of a carbon source to the medium, suggesting that these strains could either grow on a component or impurity within the agar or were able to grow on volatile compounds present in the surrounding environment. Therefore, we also manually scored growth for all strains relative to the control plate containing no carbon. Because the magnitude of colony formation differed between carbon sources, assigning a binary growth (1) or no growth (0) value was challenging for strains that exhibited residual growth on the no-carbon control or were generally slow growing. We therefore introduced a third category of nonsignificant (NS) growth for the initial screen analysis but considered them as no growth for subsequent analyses to keep the number of false positives low (fig. S11A).

The results of the computational analysis with platescan and the manual analysis showed a very high overlap between the two methods. The highest overlap (with 94.6% of matches compared with the manually scored data) was obtained with a cutoff of 20 arbitrary units (a.u.).

The comparison of the manually and computationally scored data showed that the computational method scored growth more often than the manual analysis ($n = 485$), but most of these were scored as NS initially ($n = 332$) (fig. S12). To achieve the best accuracy possible, we applied a rigorous curation process to the screening data. First, we reinspected all cases that did not agree between the manual and computational analysis (20 a.u. cutoff) and all cases that were scored as NS manually. This revealed few false-positive and false-negative scores for both the manual and computational analysis. In the case of the computational analysis, these mistakes were due to either a slight misalignment of the grid or reflections on the plate. To account for slow-growing strains and strains that exhibited high background growth on the no-carbon control condition, we also reinspected all instances that were close to the chosen pixel cutoff (10 to 40 a.u.). For these, we also inspected the growth after 10 days and scored growth on the basis of whether there was an increase in colony density over time. This latter step reduced the number of NS values observed in the initial analysis and mainly improved scoring of the previously mentioned slow-growing *Actinobacteria* (fig. S11B). The curation procedure reduced the false-positive rate compared with the original analysis from 3.3 to 2.3%, whereas the false-negative rate increased from 5.8 to 6.5%, with most uncertainty remaining for *Methylobacterium* spp. and some members of the *Actinobacteria*.

### Plant cultivation

Plants were cultivated as described previously (*19*). In brief, *A. thaliana* Columbia (Col-0) seeds were surface sterilized (*82*) and stratified for 4 days at 4°C. *Arabidopsis* were cultivated in six-well tissue culture plates (TechnoPlasticProducts) filled with 5 ml washed and heat-sterilized calcined clay mixed with 2.5 ml half-strength Murashige & Skoog medium pH 5.8 including vitamins (½ MS, Duchefa). Seeds were placed in the center of each well. If a seedling did not germinate, a different seedling was transplanted from a separate plate after 7 days. Starting at day 7, each well was supplemented twice per week with 200 μl ½ MS, respectively. Plates were incubated in a growth chamber (Percival, CU41-L4) set to 22°C and 54% relative humidity with an 11-hour photoperiod, fitted with full-spectrum lights (Philips Master TL-D 18 W/950 Graphica) and lights emitting a higher fraction of UVA and UVB (Sylvania

Reptistar F18W/6500 K). The combined light intensity was set to 200 to 210 µmol $m^{-2}$ $s^{-1}$ for wavelength 400 to 700 nm and 4 to 5 µmol $m^{-2}$ $s^{-1}$ for wavelength 280 to 400 nm.

### Phyllosphere inoculation

Bacteria grown on R-2A+M agar plates were suspended in 10 mM $MgCl_2$ solution, and the $OD_{600}$ was adjusted to 0.2 for each strain. The final inoculation suspension had an $OD_{600}$ of 0.02 for all treatments. For single-strain inoculations, 150 µl of OD-adjusted strain suspension was added to 1.35 ml 10 mM $MgCl_2$ solution. For two- and three-strain combinations, 75 or 50 µl of each strain were added, respectively. For the seven-strain community, 100 µl of each OD-adjusted strain suspension was mixed, and then a 10-fold dilution was prepared for the final inoculum. The final suspension was mixed well, and then 10-day-old *Arabidopsis* seedlings were inoculated by slowly pipetting 50 µl over the whole seedling. A 10-fold dilution series was prepared of each inoculum and spotted on R-2A+M agar plates to enumerate total bacterial load. For strain mixes, appropriate dilutions were plated on R-2A+M agar plates to verify presence of all strains on the basis of colony morphology.

### Phyllosphere harvest and bacterial colony forming unit (CFU) enumeration

Bacterial colonization in the phyllosphere was enumerated when plants were 28 or 29 days old. The whole phyllosphere was harvested with sterile tweezers and a scalpel and was placed in a 2-ml tube containing 200 µl 100 mM phosphate buffer pH 7 and a sterile 5-mm metal ball. The weight of the tube was recorded with and without the plant for plant fresh-weight calculation. The harvested phyllosphere was subsequently crushed for 45 s at 30 Hz with a TissueLyser II (Qiagen). Phosphate buffer (600 µl) was added to the crushed plant material and was mixed thoroughly; 100 µl of this suspension were transferred to a 96-well plate to prepare a 10-fold dilution series in 100 mM phosphate buffer. The dilution series was spotted on R-2A+M agar plates. In addition, 50 µl of each $10^{-3}$- and $10^{-4}$-fold dilutions were plated on 9-cm round R-2A+M agar plates. If selective plates containing streptomycin were used (for *Sphingomonas* selection), 50 µl of dilutions $10^{-1}$ and $10^{-2}$ were plated in addition. Plates were incubated at room temperature and CFUs were counted after 1 to 3 days on dilution series and after 4 to 7 days on round plates. If a strain was not found on the lowest-available dilution, its value was set to 0.9 CFUs for this sample for further analysis. *Rhodococcus* sp. Leaf233 was selectively grown on minimal medium supplemented with maltose (5 mM) when combined with *Pseudomonas* sp. Leaf15. Colonies of Leaf233 in mixtures with *Microbacterium* sp. Leaf179 were differentiated by re-streaking colonies on minimal medium containing isoleucine. Mock treated plants ($n = 12$) were included in each plant experiment to detect any systematic contamination with external bacteria.

### Shake flask cultivations

*At*-LSPHERE isolates grown on R-2A+M medium were suspended in 4 ml 10 mM magnesium chloride solution at an approximate $OD_{600}$ of 3. Each strain was inoculated at an $OD_{600}$ of 0.025 into 100-ml baffled Erlenmeyer flasks containing 10 ml of a liquid minimal medium with the same base composition (ions and vitamins) as in the carbon source screen. This medium contained all 44 growth yielding carbon sources used in the in vitro screen (table S1) at a total concentration of 10 mM C, with each carbon source at a relative concentration corresponding to the medium composition used in the modeling. Strains were inoculated in four biological replicates in monocultures and in the pairwise and community combinations used in the plant experiments. Cultures were grown at 22°C with shaking at 200 rpm for 60 hours to allow strains to reach stationary phase. Cell numbers in each culture were enumerated through a 10-fold dilution series spotted on R-2A+M agar plates.

### Data analysis

Carbon source–utilization data, plant colonization experiments, and shake flask experiments were analyzed with R 4.04 in RStudio. For the carbon source screen, false-positive and false-negative rates were calculated on the basis of 56 isolates that were screened twice (table S7). The strain phylogeny was based on full-length 16*S* ribosomal RNA gene sequences extracted from the genome sequence of each strain as described previously (*18*). Strains that did not grow on any carbon source were excluded from all further analyses. The Manhattan distances of all strains based on carbon source utilization were calculated with the *vegdist* function of the vegan 2.5 package (*83*). Hierarchical clustering was conducted with the *hclust* function with Ward method (*ward.D2*). Principal coordinate analysis was performed with the *cmdscale* function in MATLAB R2021a. Bacterial colonization data was $\log_{10}$ transformed to calculate the median colonization and statistical significance on the basis of the Wilcoxon rank sum test. *P*-values were corrected for multiple testing with the Holm method. Effect sizes for the shake flask experiments were calculated with the *cohens_d* function (var.equal = FALSE, hedges.correction = TRUE) in the *rstatix* package (*84*). $\text{Log}_2$-fold changes (combination/monoassociation) were calculated on the basis of the median absolute abundance.

Data was visualized in R with the package ggplot2 (*85*) and was further annotated with Adobe Illustrator. For data analysis and visualization, the following R packages were used: tidyverse (*86*), ape 5.4-1 (*87*), ggridges 0.5.3 (*88*), shades1.4 (*89*), phylloR (*18*), and dendextend 1.14 (*90*).

### Generation and curation of genome-scale models

The tool CarveMe (*58*) was first used to generate draft metabolic models based on the assembled genomes of each strain in the *At*-LSPHERE collection [(*53*), BioProjects PRJNA297956 and PRJEB47672]. The quality of these genomes was assessed using the tool CheckM (*91*) (table S8), which reported a mean completeness of 0.993 ± 0.007 and a mean contamination of 0.005 ± 0.007 (mean ± SD, [0,1]). We compared the predicted growth capabilities of each draft model against data from our in vitro carbon source screen and performed the following steps on each draft model: First, we used the tool NICEgame [(*62*), github.com/EPFL-LCSB/NICEgame] to generate sets of new biosynthetic and transport reactions [drawn from the BiGG database (*92*)] that would allow the model to produce biomass from each growth-supporting carbon source identified in our in vitro screen. We used this process, which relies on known reaction thermodynamic constraints (*93*, *94*), to produce at most 10 alternative sets of reactions for each growth-supporting carbon source. We separately integrated each set of reactions into the draft model and tested whether it could produce biomass on a simulated minimal medium (table S6) supplemented with each of the 45 carbon sources used in the in vitro screen. We also tested the growth of the model when combinations of reaction sets from up to three carbon sources were integrated, selecting the most accurate model (i.e., the most representative with the fewest false positives) as measured by Matthews correlation coefficient $[(TP \times TN - FP \times FN)/\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}]$ for further curation. Most models reached this stage having had sets of reactions from only one (111 models) or two (82 models) growth-supporting carbon sources integrated, as adding reactions that enable growth on one resource can resolve additional gaps that enable growth on other resources. The carbon sources most often used for gap-filling were maltose, succinate, gluconate, and xylose (being used to gap-fill 40, 26, 23, and 21 models, respectively). To correct for remaining false negatives (i.e., the model did not grow on a carbon source that supported growth in vitro), we further added reactions from a different model in our collection that did recapitulate growth on the relevant carbon source. A maximum false-positive rate of either 10 carbon sources or half of all true positives for the model, whichever was smaller, was set to avoid integrating an excessive number of new reactions. False positives for a given carbon source were corrected by removing internal reactions relevant to its metabolism or by restricting the relevant transport reaction when this was not

successful. Model- and carbon source–specific accuracies are summarized in fig. S5, and reactions added to the models as part of the curation are enumerated in fig. S4 and table S3. We identified no relationship between model accuracy and either the size (fig. S6B) or the completeness of the underlying genome (fig. S13). A final quality control was performed on each model (*59*), consisting of testing for mass and charge balance; performing a leak test to ensure no metabolites could be produced from nothing; standardizing the metabolite namespace; adding reaction subsystems; and adding additional gene, metabolite, and reaction identifiers when available. A final report was generated for each model with the validation tool MEMOTE [total score = 0.84 ± 0.02 (mean ± SD, $n = 224$)] (*95*), and balanced accuracies for individual models and pairwise interaction predictions were calculated as $(TPR + TNR)/2$, where *TPR* and *TNR* represent the true-positive rate and false-positive rates, respectively. All scripts for model generation and simulation, as well as the models and quality reports, are available at github.com/VorholtLab/i-At-LSPHERE.

### *Computing mono- and coculture growth*

All growth simulations were performed using the COBRA Toolbox v2.24.3 (*96*) with the CPLEX solver v12.10 (IBM) in MATLAB R2021a (MathWorks). Nonlimiting amounts ($v_{max} = 1000$ mmol gDW$^{-1}$ hour$^{-1}$) of a minimal medium composition containing ions, water, and sources of nitrogen, sulfur, and phosphorus were provided to the models, along with vitamins at $v_{max} = 0.15$ mmol gDW$^{-1}$ hour$^{-1}$ (table S6). Limiting amounts of the 45 carbon sources tested in our in vitro screen were provided at abundances intended to broadly estimate the relative availabilities of resource types on leaf surfaces (*36*, *38*, *39*, *54*) ($v_{max} = 0.15$ mmol gDW$^{-1}$ hour$^{-1}$ for sugars and organic acids, 0.075 mmol gDW$^{-1}$ hour$^{-1}$ for amino acids, and 1.5 mmol gDW$^{-1}$ hour$^{-1}$ for methanol). Model growth was simulated with biomass as the objective function and a minimal ATP maintenance flux of 0.5 mmol gDW$^{-1}$ hour$^{-1}$. Optimizations were also set to minimize all reaction fluxes to more closely simulate efficient proteome utilization and minimize metabolite cycling (*97*). For each pair or community, the growth of each strain was first simulated individually, and the resulting biomass flux values and resource uptake fluxes were recorded. Models were then merged by integrating them into a common extracellular compartment (*98*, *99*). Coupling constraints were introduced to avoid infeasible solutions in which one organism produced metabolic flux for the other without producing biomass itself (*34*, *100*). Additionally, metabolite uptake directionalities were fixed to those observed in monoculture to minimize inconsistencies in resource preferences between mono- and coculture. The sum of biomass reactions was set as the objective to be optimized, and exchange reaction $v_{max}$ values were set as equal to those of the models in monoculture, thus simulating an equal abundance of resources between the mono- and coculture conditions. The resulting biomass flux values were recorded and used to calculate interaction scores, which were defined as the $\log_2$ ratio of biomass flux in coculture to that in monoculture. Interaction scores were limited to between −5 and 5, with extreme $\log_2$-fold changes falling outside this range being set to −5 or 5.

## REFERENCES AND NOTES


1. A. Apprill, Marine Animal Microbiomes: Toward understanding host–microbiome interactions in a changing ocean. *Front. Mar. Sci.* **4**, 222 (2017). doi: 10.3389/fmars.2017.00222
2. E. A. Archie, J. Tung, Social behavior and the microbiome. *Curr. Opin. Behav. Sci.* **6**, 28–34 (2015). doi: 10.1016/j.cobeha.2015.07.008
3. J. Lloyd-Price *et al*., Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**, 61–66 (2017). doi: 10.1038/nature23889; pmid: 28953883
4. R. Knight *et al*., The microbiome and human biology. *Annu. Rev. Genomics Hum. Genet.* **18**, 65–86 (2017). doi: 10.1146/annurev-genom-083115-022438; pmid: 28375652
5. P. Trivedi, J. E. Leach, S. G. Tringe, T. Sa, B. K. Singh, Plant-microbiome interactions: From community assembly to plant health. *Nat. Rev. Microbiol.* **18**, 607–621 (2020). doi: 10.1038/s41579-020-0412-1; pmid: 32788714
6. L. S. Bittleston, M. Gralka, G. E. Leventhal, I. Mizrahi, O. X. Cordero, Context-dependent dynamics lead to the assembly of functionally distinct microbial communities. *Nat. Commun.* **11**, 1440 (2020). doi: 10.1038/s41467-020-15169-0; pmid: 32188849
7. J. E. Goldford *et al*., Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018). doi: 10.1126/science.aat1168; pmid: 30072533
8. S. Estrela *et al*., Functional attractors in microbial community assembly. *Cell Syst.* **13**, 29–42.e7 (2022). doi: 10.1016/j.cels.2021.09.011; pmid: 34653368
9. V. Dubinkina, Y. Fridman, P. P. Pandey, S. Maslov, Multistability and regime shifts in microbial communities explained by competition for essential nutrients. *eLife* **8**, e49720 (2019). doi: 10.7554/eLife.49720; pmid: 31756158
10. A. R. Pacheco, M. L. Osborne, D. Segrè, Non-additive microbial community responses to environmental complexity. *Nat. Commun.* **12**, 2365 (2021). doi: 10.1038/s41467-021-22426-3; pmid: 33888697
11. J. Friedman, L. M. Higgins, J. Gore, Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* **1**, 109 (2017). doi: 10.1038/s41559-017-0109; pmid: 28812687
12. O. S. Venturelli *et al*., Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* **14**, e8157 (2018). doi: 10.15252/msb.20178157; pmid: 29930200
13. A. L. Gould *et al*., Microbiome interactions shape host fitness. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11951–E11960 (2018). doi: 10.1073/pnas.1809349115; pmid: 30510004
14. T. Camenzind, S. Hättenschwiler, K. K. Treseder, A. Lehmann, M. C. Rillig, Nutrient limitation of soil microbial processes in tropical forests. *Ecol. Monogr.* **88**, 4–21 (2018). doi: 10.1002/ecm.1279
15. M. E. Hibbing, C. Fuqua, M. R. Parsek, S. B. Peterson, Bacterial competition: Surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **8**, 15–25 (2010). doi: 10.1038/nrmicro2259; pmid: 19946288
16. K. Vrede, M. Heldal, S. Norland, G. Bratbak, Elemental composition (C, N, P) and cell volume of exponentially growing and nutrient-limited bacterioplankton. *Appl. Environ. Microbiol.* **68**, 2965–2971 (2002). doi: 10.1128/AEM.68.6.2965-2971.2002; pmid: 12039756
17. J. Kehe *et al*., Positive interactions are common among culturable bacteria. *Sci. Adv.* **7**, eabi7159 (2021). doi: 10.1126/sciadv.abi7159; pmid: 34739314
18. C. I. Carlström *et al*., Synthetic microbiota reveal priority effects and keystone strains in the Arabidopsis phyllosphere. *Nat. Ecol. Evol.* **3**, 1445–1454 (2019). doi: 10.1038/s41559-019-0994-z; pmid: 31558832
19. M. Schafer, C. M. Vogel, M. Bortfeld-Miller, M. Mittelviefhaus, J. A. Vorholt, Mapping phyllosphere microbiota interactions in planta to establish genotype-phenotype relationships. *Nat. Microbiol.* **7**, 856–867 (2022). doi: 10.1038/s41564-022-01132-w; pmid: 35637327
20. K. R. Foster, T. Bell, Competition, not cooperation, dominates interactions among culturable microbial species. *Curr. Biol.* **22**, 1845–1850 (2012). doi: 10.1016/j.cub.2012.08.005; pmid: 22959348
21. R. Levins, *Evolution in Changing Environments: Some Theoretical Explorations* (Princeton Univ. Press, 1968).
22. S. H. Hurlbert, The Measurement of Niche Overlap and Some Relatives. *Ecology* **59**, 67–77 (1978). doi: 10.2307/1936632
23. M. Wilson, S. E. Lindow, Coexistence among epiphytic bacterial populations mediated through nutritional resource partitioning. *Appl. Environ. Microbiol.* **60**, 4468–4477 (1994). doi: 10.1128/aem.60.12.4468-4477.1994; pmid: 16349462
24. S. P. Li, J. Tan, X. Yang, C. Ma, L. Jiang, Niche and fitness differences determine invasion success and impact in laboratory bacterial communities. *ISME J.* **13**, 402–412 (2019). doi: 10.1038/s41396-018-0283-x; pmid: 30254322
25. M. Advani, G. Bunin, P. Mehta, Statistical physics of community ecology: A cavity solution to MacArthur's consumer resource model. *J. Stat. Mech.* **2018**, 033406 (2018). doi: 10.1088/1742-5468/aab04e; pmid: 30636966
26. E. Nestor, G. Toledano, J. Friedman, Interactions between culturable bacteria are predicted by individual species' growth. *mSystems* **8**, e0083622 (2023). doi: 10.1128/msystems.00836-22; pmid: 36815773
27. L. C. Birch, The meanings of competition. *Am. Nat.* **91**, 5–18 (1957). doi: 10.1086/281957
28. P. Chesson, Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000). doi: 10.1146/annurev.ecolsys.31.1.343
29. S. Freilich *et al*., Competitive and cooperative metabolic interactions in bacterial communities. *Nat. Commun.* **2**, 589 (2011). doi: 10.1038/ncomms1597; pmid: 22158444
30. M. Tikhonov, Community-level cohesion without cooperation. *eLife* **5**, e15747 (2016). doi: 10.7554/eLife.15747; pmid: 27310530
31. J. D. Orth, I. Thiele, B. Ø. O. Palsson, What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010). doi: 10.1038/nbt.1614; pmid: 20212490
32. A. Bordbar, J. M. Monk, Z. A. King, B. O. Palsson, Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* **15**, 107–120 (2014). doi: 10.1038/nrg3643; pmid: 24430943
33. N. Y. D. Ankrah *et al*., Enhancing microbiome research through genome-scale metabolic modeling. *mSystems* **6**, e0059921 (2021). doi: 10.1128/mSystems.00599-21; pmid: 34904863
34. S. Magnúsdóttir *et al*., Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* **35**, 81–89 (2017). doi: 10.1038/nbt.3703; pmid: 27893703
35. M. Kumar, B. Ji, K. Zengler, J. Nielsen, Modelling approaches for studying the microbiome. *Nat. Microbiol.* **4**, 1253–1267 (2019). doi: 10.1038/s41564-019-0491-9; pmid: 31337891
36. S. E. Lindow, M. T. Brandl, Microbiology of the phyllosphere. *Appl. Environ. Microbiol.* **69**, 1875–1883 (2003). doi: 10.1128/AEM.69.4.1875-1883.2003; pmid: 12676659
37. J. A. Vorholt, Microbial life in the phyllosphere. *Nat. Rev. Microbiol.* **10**, 828–840 (2012). doi: 10.1038/nrmicro2910; pmid: 23154261
38. J. H. J. Leveau, S. E. Lindow, Appetite of an epiphyte: Quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3446–3453 (2001). doi: 10.1073/pnas.061629598; pmid: 11248098
39. F. Ryffel *et al*., Metabolic footprint of epiphytic bacteria on *Arabidopsis thaliana* leaves. *ISME J.* **10**, 632–643 (2016). doi: 10.1038/ismej.2015.141; pmid: 26305156
40. G. Innerebner, C. Knief, J. A. Vorholt, Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. *Appl. Environ. Microbiol.* **77**, 3202–3210 (2011). doi: 10.1128/AEM.00133-11; pmid: 21421777
41. U. Ritpitakphong *et al*., The microbiome of the leaf surface of Arabidopsis protects against a fungal pathogen. *New Phytol.* **210**, 1033–1043 (2016). doi: 10.1111/nph.13808; pmid: 26725246
42. C. M. Vogel, D. B. Potthoff, M. Schäfer, N. Barandun, J. A. Vorholt, Protective role of the Arabidopsis leaf

microbiota against a bacterial pathogen. *Nat. Microbiol.* **6**, 1537–1548 (2021). doi: 10.1038/s41564-021-00997-7; pmid: 34819644
43. C. E. Morris, L. L. Kinkel, "Fifty years of phyllosphere microbiology: significant contributions to research in related fields" in *Phyllosphere Microbiology*, S. E. Lindow, E. I. Hecht-Poinar, V. J. Elliott, Eds. (APS Press, 2002), pp. 365–375.
44. J. Peñuelas, J. Terradas, The foliar microbiome. *Trends Plant Sci.* **19**, 278–280 (2014). doi: 10.1016/j.tplants.2013.12.007; pmid: 24439491
45. B. K. Singh, P. Trivedi, E. Egidi, C. A. Macdonald, M. Delgado-Baquerizo, Crop microbiome and sustainable agriculture. *Nat. Rev. Microbiol.* **18**, 601–602 (2020). doi: 10.1038/s41579-020-00446-y; pmid: 33037425
46. F. Martinelli *et al*., Advanced methods of plant disease detection. A review. *Agron. Sustain. Dev.* **35**, 1–25 (2015). doi: 10.1007/s13593-014-0246-1
47. A. R. Pacheco, J. A. Vorholt, Resolving metabolic interaction mechanisms in plant microbiomes. *Curr. Opin. Microbiol.* **74**, 102317 (2023). doi: 10.1016/j.mib.2023.102317; pmid: 37062173
48. A. M. Geller, A. Levy, "What I cannot create, I do not understand": Elucidating microbe-microbe interactions to facilitate plant microbiome engineering. *Curr. Opin. Microbiol.* **72**, 102283 (2023). doi: 10.1016/j.mib.2023.102283; pmid: 36868050
49. J. A. Vorholt, C. Vogel, C. I. Carlström, D. B. Müller, Establishing Causality: Opportunities of synthetic communities for plant microbiome research. *Cell Host Microbe* **22**, 142–155 (2017). doi: 10.1016/j.chom.2017.07.004; pmid: 28799900
50. A. S. Weiss *et al*., In vitro interaction network of a synthetic gut bacterial community. *ISME J.* **16**, 1095–1109 (2022). doi: 10.1038/s41396-021-01153-z; pmid: 34857933
51. J. J. Faith, P. P. Ahern, V. K. Ridaura, J. Cheng, J. I. Gordon, Identifying gut microbe-host phenotype relationships using combinatorial communities in gnotobiotic mice. *Sci. Transl. Med.* **6**, 220ra11 (2014). doi: 10.1126/scitranslmed.3008051; pmid: 24452263
52. J. Coker *et al*., A reproducible and tunable synthetic soil microbial community provides new insights into microbial ecology. *mSystems* **7**, e0095122 (2022). doi: 10.1128/msystems.00951-22; pmid: 36472419
53. Y. Bai *et al*., Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528**, 364–369 (2015). doi: 10.1038/nature16192; pmid: 26633631
54. H. B. Tukey Jr., The leaching of substances from plants. *Annu. Rev. Plant Physiol.* **21**, 305–324 (1970). doi: 10.1146/annurev.pp.21.060170.001513
55. M. J. Poupin, T. Ledger, R. Roselló-Móra, B. González, The Arabidopsis holobiont: A (re)source of insights to understand the amazing world of plant-microbe interactions. *Environ. Microbiome* **18**, 9 (2023). doi: 10.1186/s40793-023-00466-0; pmid: 36803555
56. P. Vaz Jauri, L. L. Kinkel, Nutrient overlap, genetic relatedness and spatial origin influence interaction-mediated shifts in inhibitory phenotype among Streptomyces spp. *FEMS Microbiol. Ecol.* **90**, 264–275 (2014). doi: 10.1111/1574-6941.12389; pmid: 25098381
57. M. Michalska-Smith *et al*., Network structure of resource use and niche overlap within the endophytic microbiome. *ISME J.* **16**, 435–446 (2022). doi: 10.1038/s41396-021-01080-z; pmid: 34413476
58. D. Machado, S. Andrejev, M. Tramontano, K. R. Patil, Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res.* **46**, 7542–7553 (2018). doi: 10.1093/nar/gky537; pmid: 30192979
59. I. Thiele, B. O. Palsson, A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010). doi: 10.1038/nprot.2009.203; pmid: 20057383
60. S. N. Mendoza, B. G. Olivier, D. Molenaar, B. Teusink, A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biol.* **20**, 158 (2019). doi: 10.1186/s13059-019-1769-1; pmid: 31391098
61. D. B. Bernstein, S. Sulheim, E. Almaas, D. Segrè, Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. *Genome Biol.* **22**, 64 (2021). doi: 10.1186/s13059-021-02289-z; pmid: 33602294
62. E. Vayena *et al*., A workflow for annotating the knowledge gaps in metabolic reconstructions using known and hypothetical reactions. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2211197119 (2022). doi: 10.1073/pnas.2211197119; pmid: 36343249
63. J. A. Valderrama *et al*., Bacterial degradation of benzoate: Cross-regulation between aerobic and anaerobic pathways. *J. Biol. Chem.* **287**, 10494–10508 (2012). doi: 10.1074/jbc.M111.309005; pmid: 22303008
64. G. Fuchs, M. Boll, J. Heider, Microbial degradation of aromatic compounds - from one strategy to four. *Nat. Rev. Microbiol.* **9**, 803–816 (2011). doi: 10.1038/nrmicro2652; pmid: 21963803
65. M. Murillo-Roos, H. S. M. Abdullah, M. Debbar, N. Ueberschaar, M. T. Agler, Cross-feeding niches among commensal leaf bacteria are shaped by the interaction of strain-level diversity and resource availability. *ISME J.* **16**, 2280–2289 (2022). doi: 10.1038/s41396-022-01271-2; pmid: 35768644
66. B. Ryback, M. Bortfeld-Miller, J. A. Vorholt, Metabolic adaptation to vitamin auxotrophy by leaf-associated bacteria. *ISME J.* **16**, 2712–2724 (2022). doi: 10.1038/s41396-022-01303-x; pmid: 35987782
67. L. Hemmerle *et al*., Dynamic character displacement among a pair of bacterial phyllosphere commensals in situ. *Nat. Commun.* **13**, 2836 (2022). doi: 10.1038/s41467-022-30469-3; pmid: 35595740
68. A. P. Goldberg *et al*., Emerging whole-cell modeling principles and methods. *Curr. Opin. Biotechnol.* **51**, 97–102 (2018). doi: 10.1016/j.copbio.2017.12.013; pmid: 29275251
69. D. B. Bernstein, B. Akkas, M. N. Price, A. P. Arkin, Critical assessment of E. coli genome-scale metabolic model with high-throughput mutant fitness data, bioRxiv 2023.01.05.522875 [Preprint] (2023), https://doi.org/10.1101/2023.01.05.522875.
70. C. M. J. Pieterse *et al*., Induced systemic resistance by beneficial microbes. *Annu. Rev. Phytopathol.* **52**, 347–375 (2014). doi: 10.1146/annurev-phyto-082712-102340; pmid: 24906124
71. K. Wang, A. Auzane, K. Overmyer, The immunity priming effect of the *Arabidopsis* phyllosphere resident yeast *Protomyces arabidopsidicola* strain C29. *Front. Microbiol.* **13**, 956018 (2022). doi: 10.3389/fmicb.2022.956018; pmid: 36118213
72. E. J. N. Helfrich *et al*., Bipartite interactions, antibiotic production and biosynthetic potential of the Arabidopsis leaf microbiome. *Nat. Microbiol.* **3**, 909–919 (2018). doi: 10.1038/s41564-018-0200-0; pmid: 30038309
73. B. A. Maier *et al*., A general non-self response as part of plant immunity. *Nat. Plants* **7**, 696–705 (2021). doi: 10.1038/s41477-021-00913-1; pmid: 34007033
74. R. O. Schlechter, M. Miebach, M. N. P. Remus-Emsermann, Driving factors of epiphytic bacterial communities: A review. *J. Adv. Res.* **19**, 57–65 (2019). doi: 10.1016/j.jare.2019.03.003; pmid: 31341670
75. T. Orevi, N. Kashtan, Life in a droplet: Microbial ecology in microscopic surface wetness. *Front. Microbiol.* **12**, 655459 (2021). doi: 10.3389/fmicb.2021.655459; pmid: 33927707
76. J. A. Cole, L. Kohler, J. Hedhli, Z. Luthey-Schulten, Spatially-resolved metabolic cooperativity within dense bacterial colonies. *BMC Syst. Biol.* **9**, 15 (2015). doi: 10.1186/s12918-015-0155-1; pmid: 25890263
77. E. Bauer, J. Zimmermann, F. Baldini, I. Thiele, C. Kaleta, BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLOS Comput. Biol.* **13**, e1005544 (2017). doi: 10.1371/journal.pcbi.1005544; pmid: 28531184
78. I. Dukovski *et al*., A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS). *Nat. Protoc.* **16**, 5030–5082 (2021). doi: 10.1038/s41596-021-00593-3; pmid: 34635859
79. L. Angeles-Martinez, V. Hatzimanikatis, Spatio-temporal modeling of the crowding conditions and metabolic variability in microbial communities. *PLOS Comput. Biol.* **17**, e1009140 (2021). doi: 10.1371/journal.pcbi.1009140; pmid: 34292935
80. R. Peyraud *et al*., Demonstration of the ethylmalonyl-CoA pathway by using 13C metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4846–4851 (2009). doi: 10.1073/pnas.0810932106; pmid: 19261854
81. M. Schäfer, Images of colony growth for 224 *Arabidopsis thaliana* bacterial strains on 46 carbon sources, Zenodo (2023); https://doi.org/10.5281/zenodo.7915606.
82. B. Schlesier, F. Bréton, H.-P. Mock, A hydroponic culture system for growing *Arabidopsis thaliana* plantlets under sterile conditions. *Plant Mol. Biol. Report.* **21**, 449–456 (2003). doi: 10.1007/BF02772594
83. J. Oksanen *et al*., vegan: Community Ecology Package, Version 2.6-4, 2022; https://cran.r-project.org/package=vegan.
84. A. Kassambara, rstatix: Pipe-Friendly Framework for Basic Statistical Tests, Version 0.7.2, 2023; cran.r-project.org/web/packages/rstatix.
85. H. Wickham, ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag, 2016); https://ggplot2.tidyverse.org.
86. H. Wickham *et al*., Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019). doi: 10.21105/joss.01686
87. E. Paradis, K. Schliep, ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019). doi: 10.1093/bioinformatics/bty633; pmid: 30016406
88. C. O. Wilke, ggridges: Ridgeline Plots in "ggplot2," R package Version 0.5.4, 2022; https://wilkelab.org/ggridges/.
89. J. Clayden, shades: Simple colour manipulation, Version 1.4.0, 2019; https://CRAN.R-project.org/package=shades.
90. T. Galili, dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31**, 3718–3720 (2015). doi: 10.1093/bioinformatics/btv428; pmid: 26209431
91. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015). doi: 10.1101/gr.186072.114; pmid: 25977477
92. Z. A. King *et al*., BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **44** (D1), D515–D522 (2016). doi: 10.1093/nar/gkv1049; pmid: 26476456
93. C. S. Henry, L. J. Broadbelt, V. Hatzimanikatis, Thermodynamics-based metabolic flux analysis. *Biophys. J.* **92**, 1792–1805 (2007). doi: 10.1529/biophysj.106.093138; pmid: 17172310
94. P. Salvy *et al*., pyTFA and matTFA: A Python package and a Matlab toolbox for Thermodynamics-based Flux Analysis. *Bioinformatics* **35**, 167–169 (2019). doi: 10.1093/bioinformatics/bty499; pmid: 30561545
95. C. Lieven *et al*., MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* **38**, 272–276 (2020). doi: 10.1038/s41587-020-0446-y; pmid: 32123384
96. L. Heirendt *et al*., Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14**, 639–702 (2019). doi: 10.1038/s41596-018-0098-2; pmid: 30787451
97. N. E. Lewis *et al*., Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* **6**, 390 (2010). doi: 10.1038/msb.2010.47; pmid: 20664636
98. N. Klitgord, D. Segrè, Environments that induce synthetic microbial ecosystems. *PLOS Comput. Biol.* **6**, e1001002 (2010). doi: 10.1371/journal.pcbi.1001002; pmid: 21124952
99. A. Heinken, I. Thiele, Anoxic conditions promote species-specific mutualism between gut microbes in silico. *Appl. Environ. Microbiol.* **81**, 4049–4061 (2015). doi: 10.1128/AEM.00101-15; pmid: 25841013
100. A. Heinken, S. Sahoo, R. M. T. Fleming, I. Thiele, Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* **4**, 28–40 (2013). doi: 10.4161/gmic.22370; pmid: 23022739
101. A. R. Pacheco, VorholtLab/i-At-LSPHERE: v1.0, Zenodo (2023); https://doi.org/10.5281/zenodo.7915630.
102. C. M. Field, MicrobiologyETHZ/platescan: v1.0, Zenodo (2023); https://doi.org/10.5281/zenodo.7898736.

## ACKNOWLEDGMENTS

We thank S. Hegi and M. Berger for experimental support; A. Şahin, O. Oftadeh, and D. Machado for helpful discussions on the metabolic modeling; and C. Pestalozzi for contributing to the visualization shown in Fig. 3B. **Funding:** This study was supported by Swiss National Science Foundation grant NRP72 (J.A.V.), a European Research Council Advanced Grant (PhyMo, no. 668991) (J.A.V.), Swiss National Science Foundation grant 200021_188623 (V.H.), NCCR Microbiomes, Swiss National Science Foundation (51NF40_180575) (V.H. and J.A.V.), and a James S. McDonnell Postdoctoral Fellowship (2020-1332) (A.R.P.). **Author contributions:** Conceptualization: M.S., A.R.P., and J.A.V. Methodology: M.S., A.R.P., C.M.F., E.V., and V.H. Investigation: M.S., A.R.P., R.K., and M.B.-M. Visualization: M.S. and A.R.P. Funding acquisition: A.R.P., V.H., and J.A.V. Project administration: J.A.V. Supervision: V.H. and J.A.V. Writing – original draft: M.S., A.R.P., and J.A.V. Writing – review and editing: M.S., A.R.P., R.K., M.B.-M., E.V., C.M.F., V.H., and J.A.V. **Competing interests:** The authors declare that they have no competing

interests. **Data and materials availability:** Materials and methods are available as supplementary materials. The collection of genome-scale models, as well as scripts for generating the models and conducting simulations are available through Zenodo (*101*). The platescan software (*102*) and plate images used for the carbon source screen (*81*) are available through Zenodo.
**License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf5121
Figs. S1 to S13
Tables S1 to S8
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

Submitted 27 October 2022; resubmitted 23 March 2023
Accepted 18 May 2023
10.1126/science.adf5121

RESEARCH ARTICLE SUMMARY

STRUCTURAL BIOLOGY

# Structural basis of gRNA stabilization and mRNA recognition in trypanosomal RNA editing

**Shiheng Liu†, Hong Wang†, Xiaorun Li, Fan Zhang, Jane K.J. Lee, Zihang Li, Clinton Yu, Jason J. Hu, Xiaojing Zhao, Takuma Suematsu, Ana L. Alvarez-Cabrera, Qiushi Liu, Liye Zhang, Lan Huang, Inna Aphasizheva, Ruslan Aphasizhev*, Z. Hong Zhou***

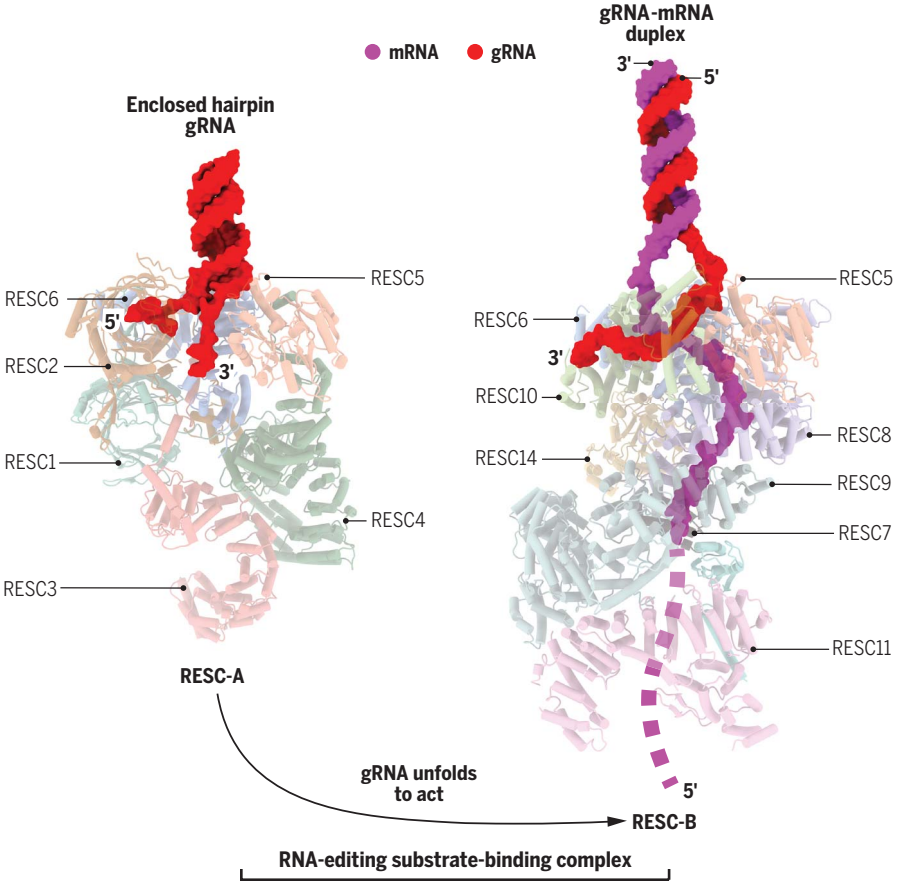
**INTRODUCTION:** Discovered in the mitochondrion of *Trypanosoma brucei*, a protozoan parasite that causes African sleeping sickness, RNA editing denotes a spectrum of phylogenetically widespread and often mechanistically unrelated molecular processes that change RNA sequence. In trypanosomal RNA editing, gRNAs direct massive recoding of cryptic mitochondrial transcripts to generate mRNAs. Assembled into the partially defined editosome, two principal ribonucleoprotein particles execute an unconventional method of genetic information transfer from gRNA to mRNA. The RNA-editing substrate-binding complex (RESC) stabilizes gRNAs and engages mRNAs. The RNA-editing catalytic complex (RECC) fulfills gRNA-programmed mRNA cleavage, uridine insertion or deletion, and religation reactions.

**RATIONALE:** Mitochondrial gRNA forms an imperfect duplex with mRNA precursor in which secondary structure defines multiple editing sites. To reveal gRNA stabilization and mRNA recognition mechanisms, we have determined atomic structures of three states of the RESC using cryogenic electron microscopy and characterized individual subunits' RNA-binding specificity.

**RESULTS:** Biochemical studies defined the RESC as a heterogenous assembly of ~18 proteins that enclose gRNAs and mRNAs. Prior work demonstrated that the RESC1/2 dimer stabilizes gRNAs and most other proteins bind mRNAs. The structure of the six-member RESC-A shows how RESC2 RNA triphosphatase pseudoenzyme engulfs the triphosphorylated gRNA's 5′ end, and how the RESC5/6 dimer fastens the 3′ end. These contacts promote gRNA folding into a "hairpin-like" conformation and shield both termini from nucleases. The 10-polypeptide RESC-B structure suggests that a remodeling event recovers gRNA from the RESC-A "storage" mode and transitions the single-stranded molecule into mRNA proximity. In the process, the gRNA's 5′ end is ejected from the RESC2 triphosphate binding tunnel but the 3′ end remains wedged between RESC5 and RESC6, which are the only proteins shared between RESC-A and RESC-B. All RESC-B subunits, including RESC5/6, contact mRNA along a ~20-nucleotide segment. However, gRNA and mRNA do not interact within RESC-B boundaries.

A typical gRNA starts with an "anchor" fully complementary to the mRNA target; the adjacent "guiding" part pairs with mRNA sparsely and creates editing sites typified by single-stranded bulges and loops. A few nucleotides separate this "information-rich" sequence from the terminal uridine tail. Mechanistically, sequestering gRNA's "information-poor" 3′ end inside RESC-B allows guiding and anchor parts to hybridize with mRNA beyond RESC-B's surface. Apparently, RESC-B proteins recruit gRNAs and mRNAs irrespective of their sequences and position the two strands in a roughly antiparallel orientation, but distant enough to prevent spurious annealing within the complex. The exposed gRNA and mRNA regions likely sample each other until productive hybridization creates a substrate for the catalytic RECC complex.

**CONCLUSION:** The architectures of distinct RESC states reveal a diversity of protein folds that have been co-opted into the RNA editing machinery. We have discovered how common gRNA elements function in stabilizing this short molecule and engaging mRNA. Our results show that gRNA-mRNA recognition emanates from ribonucleoprotein complex remodeling rather than initiating base pairing of the anchor sequence with the mRNA target. Last, structural information on essential functional features, such as the RESC2 triphosphate binding site, may facilitate development of new trypanocides. ■




**Structures of RESCs.** (Left) RESC-A sequesters gRNA termini, promoting hairpin formation and blocking mRNA access. (Right) RESC-A conversion into RESC-B unfolds gRNA and allows mRNA recognition, likely exposing editing sites to RECC-embedded enzymes.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: ruslana@bu.edu (R.A.); hong.zhou@ucla.edu (Z.H.Z.)
†These authors contributed equally to this work.
Cite this article as S. Liu *et al.*, *Science* **381**, eadg4725 (2023). DOI: 10.1126/science.adg4725

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adg4725

RESEARCH ARTICLE

STRUCTURAL BIOLOGY

# Structural basis of gRNA stabilization and mRNA recognition in trypanosomal RNA editing

Shiheng Liu[1,2]†, Hong Wang[3]†, Xiaorun Li[1,2], Fan Zhang[4], Jane K.J. Lee[1,2], Zihang Li[1,2], Clinton Yu[5], Jason J. Hu[1,2], Xiaojing Zhao[4], Takuma Suematsu[3], Ana L. Alvarez-Cabrera[1,2], Qiushi Liu[3], Liye Zhang[4,6], Lan Huang[5], Inna Aphasizheva[3], Ruslan Aphasizhev[3,7]*, Z. Hong Zhou[1,2]*

In *Trypanosoma brucei*, the editosome, composed of RNA-editing substrate-binding complex (RESC) and RNA-editing catalytic complex (RECC), orchestrates guide RNA (gRNA)–programmed editing to recode cryptic mitochondrial transcripts into messenger RNAs (mRNAs). The mechanism of information transfer from gRNA to mRNA is unclear owing to a lack of high-resolution structures for these complexes. With cryo–electron microscopy and functional studies, we have captured gRNA-stabilizing RESC-A and gRNA-mRNA–binding RESC-B and RESC-C particles. RESC-A sequesters gRNA termini, thus promoting hairpin formation and blocking mRNA access. The conversion of RESC-A into RESC-B or -C unfolds gRNA and allows mRNA selection. The ensuing gRNA-mRNA duplex protrudes from RESC-B, likely exposing editing sites to RECC-catalyzed cleavage, uridine insertion or deletion, and ligation. Our work reveals a remodeling event facilitating gRNA-mRNA hybridization and assembly of a macromolecular substrate for the editosome's catalytic modality.

Kinetoplastids are a group of flagellated protozoans that infect humans and livestock in some of the most impoverished regions of the world. *Trypanosoma brucei* spp. cause African human and animal trypanosomiasis, a substantial health threat and economic burden, respectively, in sub-Saharan Africa. The toxicity and complex regimens of existing treatments call for identification of potential parasite-specific drug targets, among which the distinct organellar RNA-editing complexes stand prominent. In trypanosomes, RNA editing restores the protein coding capacity of mitochondrial cryptogenes by massive post-transcriptional insertions and deletions of uridines (uridine, U) (*1*, *2*). Short noncoding RNAs bind to and program sequence changes in the mRNA. Astutely termed guide RNAs (gRNAs) are ~50–nucleotide (nt) molecules that feature a triphosphorylated 5′ end, followed by the "anchor" region complementary to mRNA target, the "guiding" part, and often an uridylated 3′ end (*3*–*5*). In massively edited (pan-edited) mRNAs, overlapping gRNAs sequentially bind to mRNA as recoding progresses from the 3′ to the 5′ end (*6*). Discoveries of RNA editing and gRNAs in trypanosomes have laid the foundation for developing modern RNA-directed genome- and transcriptome-altering technologies.

The partially defined *T. brucei* editosome includes RNA-editing substrate-binding complex (RESC) and RNA-editing catalytic complex (RECC) (*7*–*9*). Molecular studies have implicated homologous binding partners RESC1 and RESC2, previously named GRBC1/2 (*10*) or GAP2/1 (*11*), in gRNA stabilization and demonstrated their presence in heterogeneous (0.3 to 1.2 MDa) ribonucleoproteins (RNPs) involving up to 18 polypeptides [RESC protein nomenclature is available in (*8*)] (*10*–*12*). The RNPs that contain the RESC1/2 dimer incorporate editing substrates (gRNAs and pre-edited mRNAs) and products (partially and fully edited mRNAs) (*7*, *12*–*14*). Conversely, the RECC complex encloses enzymes performing gRNA-programmed mRNA cleavage, U insertions or deletions, and religation (*15*–*17*).

Efforts to understand editing mechanisms at the structural level remain limited to crystallographic studies of recombinantly produced individual enzymes (*18*–*20*) and factors (*21*–*23*). Previous electron microscopy (EM) attempts yielded negative-stain structures of the catalytic editing complex at ~30-Å resolution (*24*, *25*). In this study, we combined biochemical, mass spectrometry, in vivo RNA profiling, and cryoID (*26*) approaches to capture three states of the RESC and reveal their atomic details by single-particle cryo–electron microscopy (cryo-EM).

## Results

### *Capturing three states of RESC: RESC-A, RESC-B, and RESC-C*

To produce isolates amenable to cryo-EM reconstitutions at near-atomic resolution, we initially purified endogenous RESC2 (Fig. 1A and fig. S1) and separated the associated complexes by glycerol gradient and native polyacrylamide gel electrophoresis (PAGE) to resolve the predominant ~280 kDa and less abundant ~0.5 to 1.2 MDa complexes (Fig. 1B). Although the RESC1/2 dimer binds gRNAs in vitro (*10*, *27*), distribution along the gradient shows that the most abundant ~280 kDa RESC1/2-containing particles are devoid of gRNAs, which are sequestered into complexes exceeding 400 kDa (Fig. 1B, bottom). To capture these larger gRNA-containing assemblies, we tested RESC2, RESC5, RESC9, and RESC14 isolates for RESC components and associated mRNA-processing complexes (fig. S2 and table S1). The RESC2-purified sample contained abundant RESC1/2 and KREH2 helicase, whereas both RESC9 and RESC14 isolates lacked RESC1 to RESC4 and RESC15 to RESC18, respectively. Compared with the others, the RESC5 complex comprised a more uniform set of RESC proteins and higher relative amount of the catalytic RECC (fig. S2). Size separation of RESC5 (Fig. 1C) demonstrated cosedimentation with RESC4 (Fig. 1D), which is predicted to interact with RESC1/2 and RESC5 (*7*, *13*). Likewise, native PAGE revealed RESC5 presence in distinct complexes (Fig. 1D), of which RESC1/2 co-occupies the ~400-kDa particles and those exceeding 600 kDa (Fig. 1E). Informed by these findings, we selected RESC5 for cryo-EM analysis. To mitigate sample heterogeneity, mitochondrial lysate was treated with ribonucleases (RNases) before purification; this step reduced copurification of RESC12/12A mRNA-binding proteins (*28*, *29*) and mRNA editing (RECC and KREH2C) and processing (KPAC and PPsome) complexes (*8*) (Fig. 1F and table S2).

The RNase-treated RESC5 sample was evaluated by means of negative-stain EM and cryo-EM. The 2D averages of the cryo-EM particles revealed multiple species with miscellaneous shapes and dimensions (fig. S3, A and B). By combining these 2D classes for in-depth 3D analysis (fig. S4), we obtained a series of 3D structures, among which three were identified as RESC protein–containing complexes through the cryoID approach (*26*), and their atomic models were built. We designated them as RESC-A (3.7 Å), RESC-B (3.4 Å), and RESC-C (3.3 Å) (Fig. 1G, figs. S9 and S10, table S3, and movie S1). RESC-A contains RESC1 to RESC6, including the gRNA-stabilizing heterodimer RESC1/2. RESC-B incorporates RESC5 to RESC11, RESC13, and RESC14, thus sharing the RESC5/6 heterodimer with RESC-A and possessing the mRNA-interacting RESC13 (*28*, *29*). The smallest complex, RESC-C, consists of RESC5 to RESC8, RESC10, and RESC14; it may represent a RESC-B assembly intermediate or reflect RESC-B remodeling during editing (Fig. 1G and fig. S8). Most RESC components are composed of repetitive arrays of short amphiphilic α helices

[1]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA, USA. [2]California NanoSystems Institute, University of California, Los Angeles, CA, USA. [3]Department of Molecular and Cell Biology, Boston University Medical Campus, Boston, MA, USA. [4]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [5]Department of Physiology and Biophysics, University of California, Irvine, CA, USA. [6]Shanghai Clinical Research and Trial Center, Shanghai 201210, China. [7]Department of Biochemistry, Boston University Medical Campus, Boston, MA, USA.
*Corresponding author. Email: ruslana@bu.edu (R.A.); hong.zhou@ucla.edu (Z.H.Z.)
†These authors contributed equally to this work.

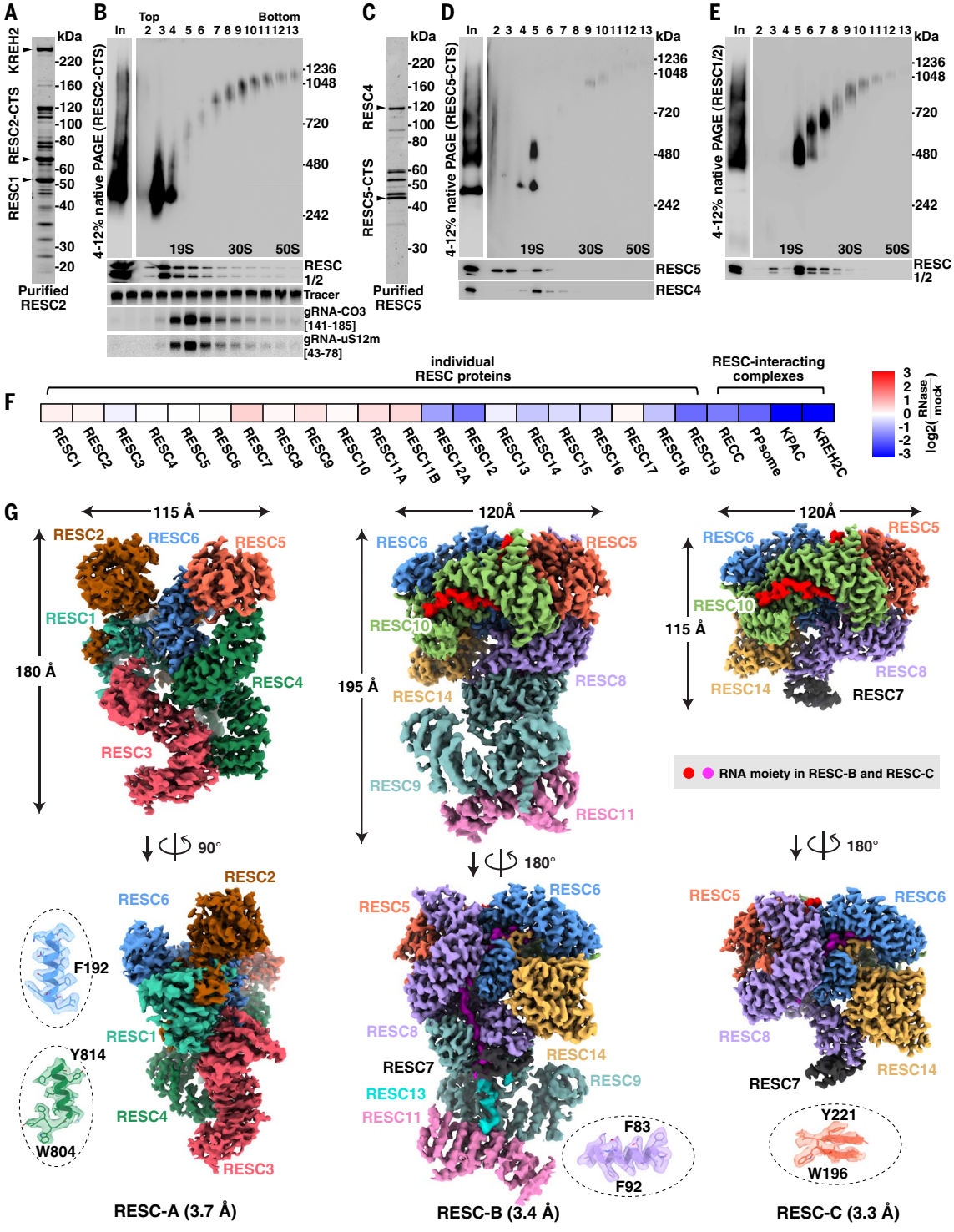



**Fig. 1. Overall architecture of RESC-A, -B, and -C complexes from *T. brucei*.** (**A**) Tandem affinity–purified CTS-tagged RESC2 separated by means of SDS-PAGE and stained with fluorescent dye. Arrows indicate gRNA-binding proteins RESC1 and RESC2 and DEAH/RHA RNA helicase KREH2. (**B**) RESC2-CTS isolate separated by means of native PAGE (left) or glycerol gradient and then native PAGE (right), and visualized by immunoblot with a tag-specific antibody. Sedimentation rate values (*S*) of thyroglobulin, and small and large bacterial ribosomal subunits align with fractions probed for RESC1/2 (top; immunoblot with antibodies recognizing both proteins) and gRNAs CO3 (141 to 185) and uS12m (43 to 78]) (bottom; northern blot). Synthetic radiolabeled tracer shows uniformity of RNA isolation from gradient fractions (middle). (**C**) Tandem affinity–purified CTS-tagged RESC5 separated by means of SDS-PAGE and stained with fluorescent dye. (**D**) RESC5-CTS isolate separated by means of native PAGE (left), or by glycerol gradient and then native PAGE (right) and visualized with tag-specific antibody. RESC4 and RESC5 were detected in each fraction by immunoblot. (**E**) Membranes shown in (D) probed with antibodies against RESC1/2. (**F**) Relative abundances of RESC proteins and RNA-editing catalytic (RECC), mRNA stabilizing (PPsome), polyadenylation (KPAC), and KREH2 RNA helicase (KREH2C) complexes in RESC5 isolates from mock- and RNase-treated extracts. Heatmap represents log2-fold change values. RNase treatment of the extract prevented RESC5 co-isolation with RECC, PPsome, KPAC, and KREH2C; RESC12 and RESC12A paralogs also declined. Supporting mass spectrometry data are provided in table S2. (**G**) Distinct views of RESC-A (left), RESC-B (middle), and RESC-C (right) cryo-EM density maps. (Insets) Models overlayed on transparent cryo-EM density. Additional overlays are provided in fig. S9.

called HEAT repeats (also known as α solenoid) (fig. S11). HEAT repeats often form extended superhelical structures and mediate protein-protein interactions (*30*, *31*), as observed here. Apparent molecular masses of RESC5 complexes agree with values calculated from structural models; RESC-A (~380 kDa) and RESC-B (~600 kDa) likely populate the upper band, with the lower band representing RESC-C (~270 kDa) (Fig. 1D and fig. S8).

Depletion of RESC12 and RESC12A from RESC5 isolates through RNase treatment (Fig. 1F) led us to inquire whether these paralogs that share 77% protein sequence identity and bind to pre-edited mRNA (*28*, *29*) connect with

RESC by means of an RNA component (Fig. 1F). For RESC12 and RES12A isolations without RNase treatment, paralog-specific peptide spectra counting showed reciprocal RESC12 and RESC12A relative abundances and their similar levels in RESC5 isolate (table S4). Overall, we conclude that RESC-A and RESC-B interact with KREH2 helicase and pre-edited mRNA binding factors, respectively, through RNA. As detailed below, the atomic models of RESC-A, RESC-B, and RESC-C reveal interactions among their subunits, the mechanisms of gRNA stabilization, and gRNA-mRNA recognition. Our analyses also suggest the mechanism by which a ribonucleoprotein substrate for the catalytic RECC assembles.

### *Structure of RESC-A*

RESC-A contains six proteins and has dimensions of 180 by 115 by 95 Å (Figs. 1G and 2A). The presence of the RNA triphosphate tunnel metalloenzyme (TTM) domain (*32*) typifies RESC1 and RESC2; the latter also has an N-terminal armadillo (ARM) domain. RESC5 and RESC6 consist of β-propeller and superhelical HEAT domains, respectively. RESC3 and RESC4 are characterized by superhelical HEAT domains flanked by an N-terminal anchor and a C-terminal clip domain, respectively (fig. S11). RESC1 and RESC2 share a protein sequence identity of 31% and form a heterodimer (Fig. 2B and fig. S13B). A hairpin loop (amino acids 181 to 212) protruding from RESC2's TTM runs along the crevice outside the triphosphate tunnel of RESC1, yielding extensive hydrophobic interactions and main chain–main chain hydrogen bonds (Fig. 2B, middle and right). Hydrogen bonds between RESC2's ARM and RESC1's TTM further enhance association (Fig. 2B, top left). Extensive RESC1/2 contacts likely underlie their interdependent persistence in the cell (*11*). The RESC1/2 heterodimer occupies the superhelical inner surface of RESC6's HEAT by means of the extended loop of RESC2's ARM, forming hydrogen bonds (Fig. 2C). RESC5 associates with RESC6 through hydrogen bonds and hydrophobic contacts (Fig. 2D), generating a positively charged crevice between RESC5 and RESC6 that is potentially receptive to RNA binding (Fig. 2E). RESC3 and RESC4 stabilize the orientations of RESC1/2 and RESC5/6 heterodimers (Fig. 2A and fig. S12, A and B).

### *Guide RNA interactions in RESC-A*

The RESC-A structure described above was determined in the RNase-treated isolate (Figs. 1F and 2A, and figs. S4 and S5). Because molecular studies identified RESC1/2 as the key gRNA-stabilizing element (*10*), we next isolated RESC5 from mock-treated lysate and determined RESC-A structure with gRNA at 3.7-Å resolution (Fig. 3A, figs. S6 and S7, and movie S2). The overall structures are similar (Fig. 3B, left), except that disordered loops close to gRNA in RESC1 and RESC2 become organized and visible in the presence of gRNA (Fig. 3B, inset). The triphosphate binding tunnel of RESC2 engulfs the gRNA's 5′ end while the 3′ end docks into the crevice formed by RESC5 and RESC6 (Fig. 3, A and C). Specifically, the gRNA's 5′ triphosphate and conserved transcription initiation sequence contact the triphosphate binding tunnel, whereby the "anchor" and "guiding" regions hybridize into an ~18–base pair (bp) stem (Fig. 3D) reminiscent of gRNA folding in solution (*33*, *34*). The locking of anchor and guiding regions into a hairpin suggests that gRNA must undergo a drastic conformational change to recognize mRNA. In addition to the 5 nt at the 3′ end lining up the crevice between RESC5/6, a total of 47 nt were modeled for gRNA in RESC-A (Fig. 3D). Accounting for the unstructured loop region, the RESC-bound gRNA matches the median 49-nt length of encoded gRNA sequences (*35*). Our findings imply that a hairpin conformation and 3′-end shielding by RESC5/6 render gRNA refractory to 3′-to-5′ degradation (*4*, *36*, *37*). We conclude that RESC-A, rather than solely the RESC1/2 heterodimer (*10*, *11*, *27*), is the gRNA-stabilizing particle.

gRNA is the only class of mitochondrial transcripts that retains 5′ triphosphate characteristic of transcription start site (*3*, *38*), whereas the nonencoded 3′ U tail is added by the processome after precursor trimming (*4*, *5*). To investigate the functional importance of the triphosphate binding by RESC2, we superimposed RESC1 and RESC2 TTM domains with Cet1 mRNA triphosphatase from *Saccharomyces cerevisiae* (*32*). Structural comparisons and sequence alignments of *T. brucei* and related parasite *Leishmania major* proteins show substitutions of two glutamic acids in the catalytic metal binding triad (E305-E307-E496; Cet1 numbering) that typify RESC1 and RESC2 as hydrolytically inactive pseudoenzymes (Fig. 3, E to G). The lack of phosphatase activities in RESC5 isolate and reconstituted RESC1/2 heterodimer support this conclusion (fig. S13). However, in the gRNA-occupied RESC2 TTM tunnel, positively charged residues (K311, R402, and R424; RESC2 numbering) form hydrogen bonds with the 5′ triphosphate moiety (Fig. 3, D and F). In RESC1, substitutions of potential triphosphate binding positions (R393 and K456; Cet1 numbering) by acidic or polar side chains (Fig. 3G) likely explain the lack of gRNA occupancy (Fig. 3A). To test the essentiality of triphosphate contacts, we constructed cell lines for inducible overexpression of RESC2 mutants and tested their growth kinetics. Replacing E240A/N242A, residues in positions structurally homologous to Cet1 catalytic metal-binding side chains E305/E307, exerted no growth inhibition, which is consistent with the lack of a $Mg^{2+}$-binding site in RESC2 (Fig. 3, F and G). By comparison, mutations of triphosphate-binding amino acids caused growth-inhibition phenotypes ranging from mild (K311A) to moderate (R402A/K406A) to severe (R424A) (Fig. 3H). As a control, R408A mutation in RESC1 (corresponding to R424A in RESC2) produced no discernible growth defect (fig. S14).

The long-debated U-tail function (*3*, *5*, *39*, *40*) compelled a closer scrutiny of the 3′-end binding in the RESC5/6 crevice (Figs. 2E and 3, C and D). Ultraviolet (UV) irradiation of live parasites and sequencing of RNAs cross-linked to RESC1/2 and RESC5/6 (table S5) demonstrate that RESC-A shields the entire gRNA and indicate that RESC5 binds longer gRNA fragments than RESC2 (Fig. 3I). The difference emanates from RESC5-bound gRNAs possessing extended U tails (Fig. 3J). Consistent with uridylation patterns of predicted gRNAs versus gRNA-like molecules encoded by pseudo-gRNA genes (*35*), we found that RESC5 functionally selects gRNAs with longer U tails (Fig. 3K). Conversely, the lack of U-enriched motifs in RESC5-bound fragments stipulates a "molecular ruler" mechanism of gRNA selection: The length, rather than the sequence of nonencoded additions, delivers the 3′ end into the crevice between RESC5 and RESC6 (Fig. 3C). Taken together, our structural and functional data show that RESC-A selects gRNA by recognizing the 5′ triphosphate and accommodating a defined length and the capacity to fold into a hairpin-like conformation.

### *mRNA recognition in RESC-B*

We also determined structures of RESC-B from RNase- and mock-treated RESC5 isolates. Although the latter has extended RNA densities and additional protein components, their overall structures are similar. The following description refers to their common features unless otherwise stated: RESC-B shares RESC5/6 heterodimer with RESC-A and contains additional components RESC7 to RESC14 [Figs. 1G (RNase-treated) and 4A (mock-treated) and movie S3]. Multiple helix-loop-helix structural motifs populate subunits RESC7-12 (fig. S11), with RESC8 to RESC12 forming extended superhelical HEAT domains. RESC13's N-terminal RGG and C-terminal RRM motifs, both implicated in mRNA binding (*41*), are present in the structure of mock-treated RESC-B; only the RGG domain is visible in RNase-treated RESC-B. The RESC14 phytanoyl-CoA dioxygenase pseudoenzyme's phyH domain is wedged between RESC6, RESC8, and RESC10 (Fig. 4A and fig. S12, C and D).

Two RNA strands are embedded in RESC-B. The shorter fragment contacts RESC5, RESC6, and RESC10 (Fig. 4, A and B, and fig. S12C), with nucleotides −16 to −12 (we designate position −1 as the 3′ end of visible RNA) being stacked between RESC5 and RESC6 (Fig. 4D), which mimics gRNA binding in RESC-A (Fig.

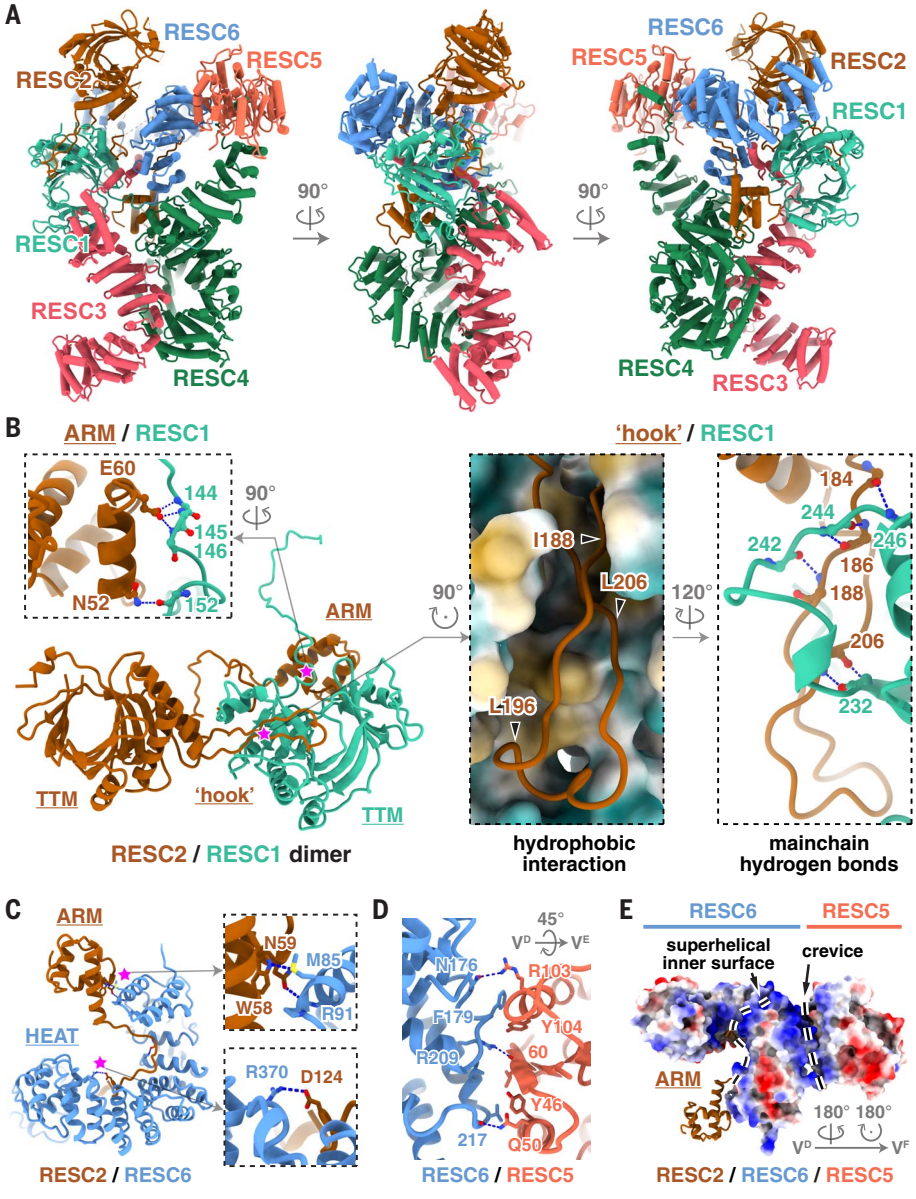



**Fig. 2. Structure of RESC-A protein complex.** (**A**) Perpendicular views of the overall RESC-A assembly. (**B**) Magenta stars point to close-up views of interacting regions in the RESC1/2 heterodimer, detailing interactions between the ARM domains of RESC2 and RESC1 (top left) and between the RESC2 "hook" domain and RESC1 (right). Yellow patches color hydrophobic interactions with RESC2, presented as a strand, and RESC1, shown as a surface. (**C**) The C-terminal loop of the RESC2 ARM domain engages the inner surface of the RESC6 HEAT domain superhelix. Magenta stars point to expanded views of amino acid residues in the contacting regions. (**D**) Side-chain interactions between RESC5 and RESC6. (**E**) Electrostatic surface potential of the RESC5/6 heterodimer.

3, C and D). Therefore, this nucleotide strand was tentatively designated gRNA. In contrast to RESC-A, the 5′ gRNA nucleotides exiting from RESC5/6 are not visible in RNase-treated RESC-B, whereas the 3′ nucleotides extend further and interact with RESC10 (Fig. 4, A and D). Comparison of RESC-A and RESC-B suggests gRNA transition from a hairpin in RESC-A (Fig. 3, B to D) into single-stranded conformation in RESC-B (Fig. 4, A, B, and D). The enduring gRNA attachment to RESC5/6 during remodeling renders the stability of this interaction as a quality checkpoint for tunneling of properly 3′-processed molecules into the editing cascade.

RNA strand modeling and in vivo RNA-protein cross-linking data suggest that mRNA fragments account for the longer nucleotide segment traversing the entire RESC-B and interacting with all proteins (Fig. 4, A, C, E, H, and I, and fig. S12, D to F). The unmodeled 5′ moiety runs through the interval between the RGG and RRM domain of RESC13 and further associates with the inner surface of RESC11's HEAT superhelix (Fig. 4, F and G). Heading toward the 3′ end, the mRNA is sequentially shielded by the superhelical inner surface of the HEAT repeats of RESC9, RESC8, and RESC6. This shielding is consistent with residual RNA moiety in the RNase-treated RESC-B (Fig. 1G, middle). In RESC-B, the mRNA's 3′ region is positioned in proximity (~30 Å) to the 5′ end of the single-stranded gRNA fragment unwound by RESC5, RESC6, and RESC10 (Fig. 4B). This mutual orientation potentially allows gRNA and mRNA to hybridize beyond the RESC-B surface.

### *gRNA-mRNA duplex scaffolding by RESC-B*

Our cryo-EM structures show gRNA binding by RESC1/2 and RESC5/6 in RESC-A, and a continued gRNA engagement with RESC5/6 and additional interactions with RESC10 in RESC-B, all subunits of which also contact mRNA. However, sequence diversity of the endogenous transcripts impeded identification by cryo-EM of RNAs bound to each protein. To directly identify RNAs bound by individual RESC components, we carried out UV cross-linking of live parasites and purification of RNA-protein cross-links, RNA fragmentation, and sequencing (figs. S1 and S15A and table S5) (*42*). Consistent with our structures, RESC1 and RESC2 (RESC-A) display strong preference for gRNAs; RESC5, RESC6 (RESC-A and RESC-B), and RESC10 (RESC-B) cluster in the intermediate range commensurate with dual gRNA-mRNA binding modality; RESC-B–specific RESC7 to RESC14 predominantly cross-link to mRNAs (Fig. 4H). These data confirm the identity of the nucleotide strand held by RESC5/6/10 as gRNA (Fig. 4D) and the strand traversing the entire RESC-B as mRNA (Fig. 4A).

By forming an ~11-bp "anchor" duplex, which lengthens as U insertions and deletions extend complementarity with mRNA, gRNA initially recognizes cognate mRNA. Considering the lack of gRNA-mRNA pairing within RESC-B, we next assessed editing prevalence in mRNA fragments cross-linked to RESC-B proteins. The ratios of fully edited sequences to all reads derived from pan-edited mRNAs meeting coverage thresholds, uS12m (Fig. 4I), A6, and CO3, and from moderately edited CYB mRNA (fig. S15B and table S5), demonstrate that RESC12/12A paralogs and RESC13 display strongest preference for pre-edited mRNAs. The extent of editing increases in RESC14, whose suppression spares initiation but blocks editing progression (*43*, *44*). As expected from our structures, edited sequences are most represented among RNAs cross-linked to gRNA-mRNA–binding RESC5, RESC6, and RESC10 (Fig. 4I). Longitudinal mapping of RESC5 and RESC12 (Fig. 4J) and other RESC-B proteins (fig. S16) in vivo occupancies of uS12m

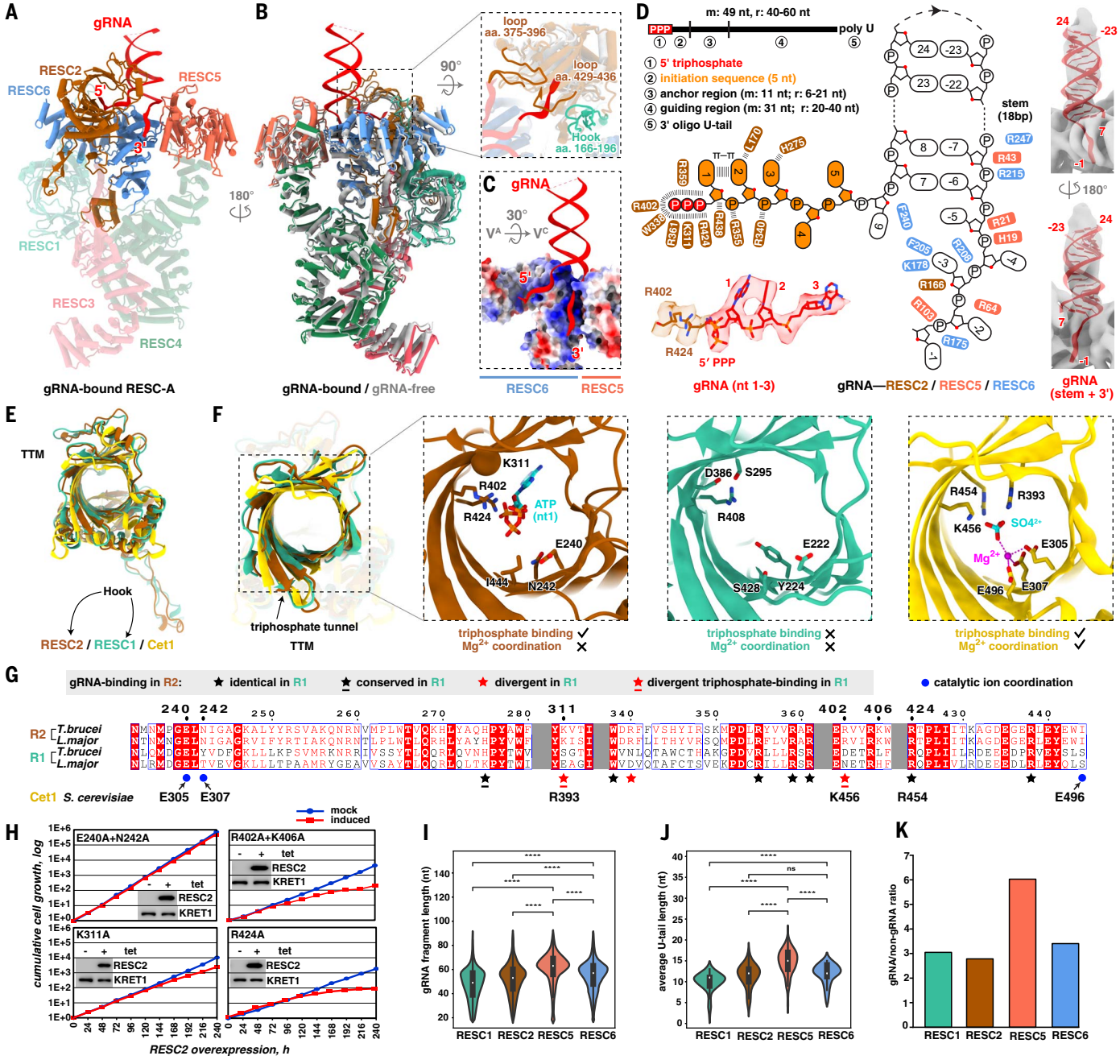



**Fig. 3. Structure of RESC-A complex with gRNA.** (**A**) The gRNA binding RESC1/2 and RESC5/6 heterodimers and gRNA are highlighted in the overall RESC-A structure. (**B**) Disordered loops in RNA-depleted RESC-A become structured and visible with bound gRNA. (**C**) The 3′ terminus of gRNA occupies the positively charged crevice formed by RESC5 and RESC6. (**D**) Interactions between gRNA and gRNA-binding proteins. An atomic model was built for the first three nucleotides, with the triphosphate group visible on the first nucleotide. The rest of gRNA moiety was traced with the main chain. The model overlayed on transparent cryo-EM density map was shown for the first nucleotides of gRNA (bottom left) and the gRNA stem and the 3′ fragment (right). *m*, median length; *r*, range. (**E**) Superimposition of triphosphate tunnel metalloenzyme (TTM) domains from RESC2 (brown), RESC1 (turquoise), and *S. cerevisiae* RNA triphosphatase Cet1 (yellow). (**F**) Focused views of RNA 5′ triphosphate-binding and catalytic metal ion–coordination sites in RESC2, RESC1, and Cet1. (**G**) Multiple sequence alignment of the TTM domains in RESC2 (R2) and RESC1 (R1). Red background depicts identical residues. Conserved residues are colored in red, and divergent residues are in black. gRNA-binding residues in RESC2 are labeled with stars; underlined red stars indicate RESC2 triphosphate–binding side chains that are divergent in RESC1. Cet1 residues structurally homologous to RESC2 positions are numbered; blue circles indicate catalytic metal-binding triad. (**H**) Growth kinetics of parasite cell lines that conditionally express mutated RESC2 proteins. (Insets) RESC2 expression in mock (−tet) cells and cells that were tetracycline-induced for 72 hours (+tet). KRET1, loading control. (**I**) Size distribution of gRNA fragments cross-linked to RESC1, RESC2, RESC5, and RESC6. ****$q$ < 0.0001, Welch's $t$ test followed by Benjamini–Hochberg (BH) false discovery rate (FDR) correction. (**J**) Distribution of average U-tail length in gRNAs fragments cross-linked to RESC1, RESC2, RESC5, and RESC6. ****$q$ < 0.0001; ns, not significant. Mann-Whitney $U$ test, followed by BH FDR correction. (**K**) Ratios of gRNAs versus minicircle-encoded gRNA-like molecules cross-linked to RESC-A subunits. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; E, Glu; K, Lys; R, Arg; N, Asn.

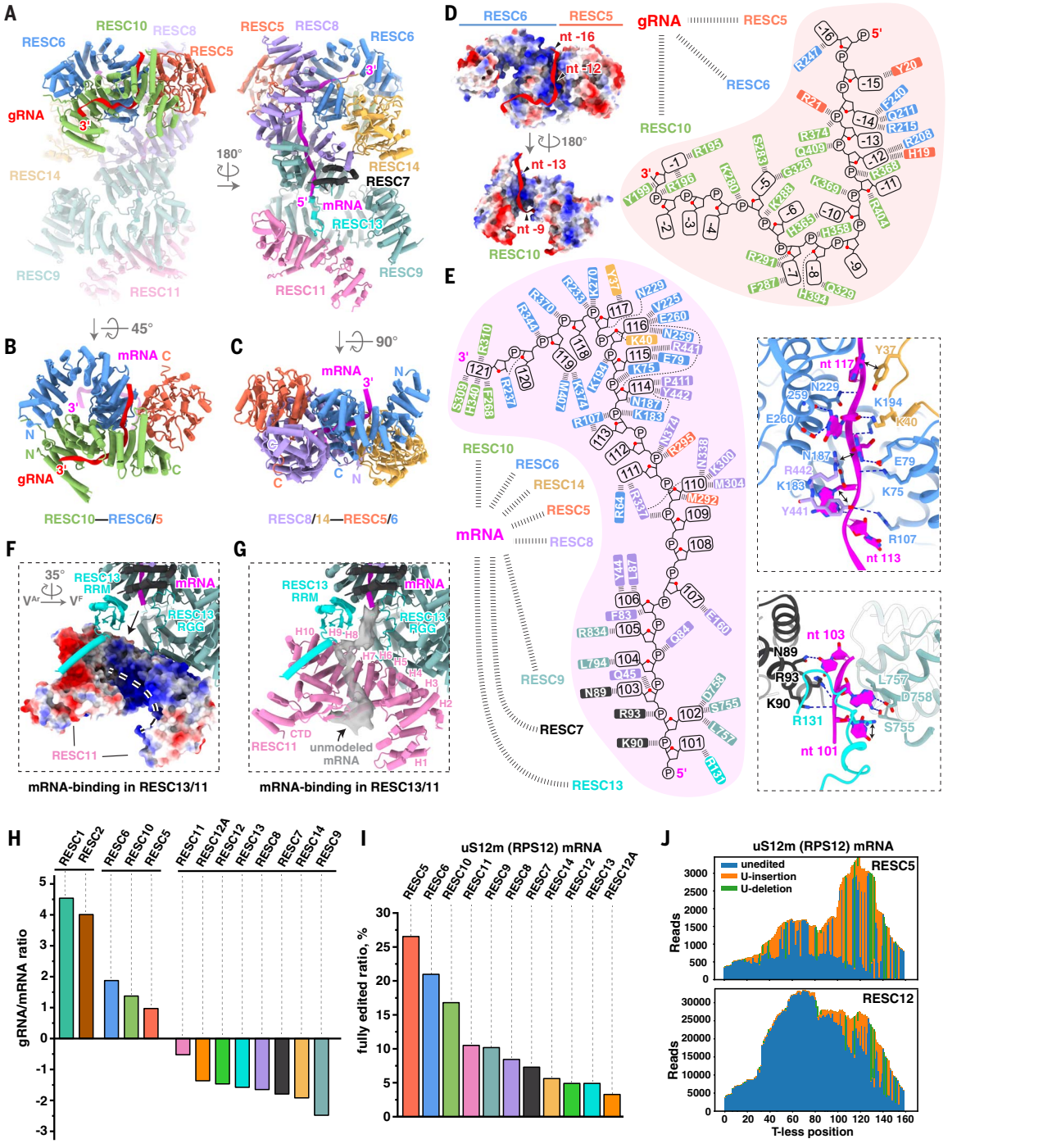



**Fig. 4. gRNA and mRNA trajectories in RESC-B complex.** (**A**) Overall RESC-B structure. (Left) gRNA-interacting proteins RESC5, RESC6, and RESC10 and gRNA moiety. (Right) mRNA-binding proteins RESC5, RESC6, RESC7, RESC8, RESC9, and RESC11 and the mRNA reconstitution. (**B**) Binding of gRNA and mRNA fragments by RESC5, RESC6, and RESC10. (**C**) Binding of the mRNA's 3′ region by RESC5, RESC6, RESC8, and RESC14. (**D**) Electrostatic interactions between gRNA and RESC5, RESC6, and RESC10. Blue surfaces show positively charged grooves. (**E**) Interactions between mRNA and RESC-B residues. Close-up views show the detailed interaction of RESC7, RESC9, and RESC13 with mRNA positions 101 to 103, and those of RESC6, RESC8, and RESC14 with positions 113 to 117. (**F** and **G**) Potential architecture of mRNA 5′-end binding site shaped by RESC11 and RESC13. (F) RESC11 is shown as electrostatic surface potential, whereas RESC13 is shown as ribbon. (G) RESC11/RESC13 are shown as ribbon, whereas unmodeled mRNA is shown as transparent gray density. Double-dashed line indicates positively charged surface area. (**H**) In vivo gRNA-mRNA binding preferences of RESC proteins. Reads were mapped to annotated gRNAs (*35*) and to 12 edited mRNA genes in all editing states. Read-count ratios are log2-transformed and centered by the average of all RESC proteins. (**I**) Prevalence of fully edited uS12m (RPS12) mRNA sequences detected with in vivo UV cross-linking analysis of RESC-B proteins. The ratio of fully edited to combined pre-, partially, and fully edited mRNA fragments is plotted for each RESC-B subunit. (**J**) Positional mapping of RESC5 and RESC12 in-vivo binding sites in uS12m (RPS12) mRNA. The *x* axis represents 5′-to-3′ mRNA coordinates after removing encoded Us and those inserted by editing (T-less). The *y* axis shows the total read depth for pre-edited (blue), U-insertions (orange), and U deletions (green) per nucleotide.

mRNA further corroborates their aggregate binding preferences ranging from mostly pre-edited (RESC11/12/13) to partially edited (RESC5/6/10) sequences.

A gRNA-mRNA duplex protruding from the RESC5/6/10 cluster would be expected from structural and RNA profiling data. Local 3D classification of RESC-B particles indeed shows an extended gRNA-mRNA duplex (Fig. 5 and fig. S10J). The resultant reconstitutions traced a nearly complete 43-nt gRNA exiting from RESC-B and forming a 24-bp duplex with the 3′ part of the contiguous 51-nt mRNA fragment (Fig. 5A).

In typical gRNA, the guiding part is separated from the U tail by a few nucleotides that cannot pair with mRNA (*35*). Mechanistically, sequestering gRNA's "information-poor" 3′ end by RESC5/6/10 allows guiding and anchor parts to hybridize with mRNA beyond RESC-B surface, hence, accessible to RECC (Fig. 5C). It follows that RESC5/6/10 likely retains the U tail and adjacent nonguiding nucleotides as editing events take place between the 5′ anchor and gRNA's 3′ end. Because the mRNA is cleaved at each editing site, RESC-B contacts with single-stranded gRNA 3′ and mRNA 5′ regions likely tether mRNA cleavage fragments to gRNA. As documented by "precleaved" editing assays, such tethering stimulates U insertion, U deletion, and RNA ligation activities (*45*–*47*). Collectively, our findings introduce RESC-B as the probable substrate for the catalytic RECC complex. Upon completion of editing directed by a single gRNA, RESC-B may also expose the double-stranded region formed by gRNA and fully edited mRNA to postediting transactions.

## Discussion

By using the cryoID approach, we have identified and built atomic models of three RNA-editing substrate-binding complexes—RESC-A, RESC-B, and RESC-C—from *T. brucei* mitochondrion. Information stored in minicircle and maxicircle genomes conflates during U insertion and deletion editing, whereby gRNAs produced from the former direct posttranscriptional recoding of cryptic transcripts encoded in the latter. The RESC molecular machines pairing hundreds of gRNAs with their targets are key to this broadly important process. In RESC-A composed of RESC1 to RESC6, RESC2 pseudotriphosphatase sequesters gRNA's 5′ end into the triphosphate binding tunnel. Homologous binding partner RESC1 attaches RESC2 to RESC-A, wherein RESC5 and RESC6 affix gRNA's 3′ end by selecting molecules with longer U tails. The ensuing hairpin formation and protein contacts protect gRNA against 3′-to-5′ degradation in a plausibly inactive conformation whereby the anchor hybridizes with the guiding region. Such "closed" conformation implies that gRNA remodeling precedes mRNA recognition. In RESC-B, composed of RESC5 to RESC14, we observed an mRNA fragment traversing the entire complex. An extensive hydrogen-bonding network fastens mRNA to every protein and positions its 3′ part close to a single-stranded gRNA segment unwound by interactions with RESC5, RESC6, and RESC10. Smaller RESC-C also contains gRNA and mRNA; whether this abundant particle represents a stable intermediate in the transition from RESC-A to RESC-B, or reflects RESC-B disassembly during postediting remodeling—e.g., pre-edited mRNA moving out of RESC12/12A—remains to be established.

Consistent with our atomic models, in vivo RNA-occupancy profiling discerned RESC components into preferential gRNA (RESC1/2), dual gRNA-mRNA (RESC5, RESC6, and RESC10), and pre-edited mRNA (RESC7 to RESC9 and RESC11 to RESC14) binding factors (Fig. 4H). We observed the least editing in mRNA fragments bound to RESC12 and RESC13, which are distal to gRNA in the RESC-B structure, and the most in those cross-linked to gRNA-proximal RESC5, RESC6, and RESC10 (Fig. 4I). The extent of U insertions and deletions gradually increases from the 5′ mRNA section sequestered by the RESC12/13 "entry" cluster toward the RESC5/6/10 "exit" module. Apparently, the

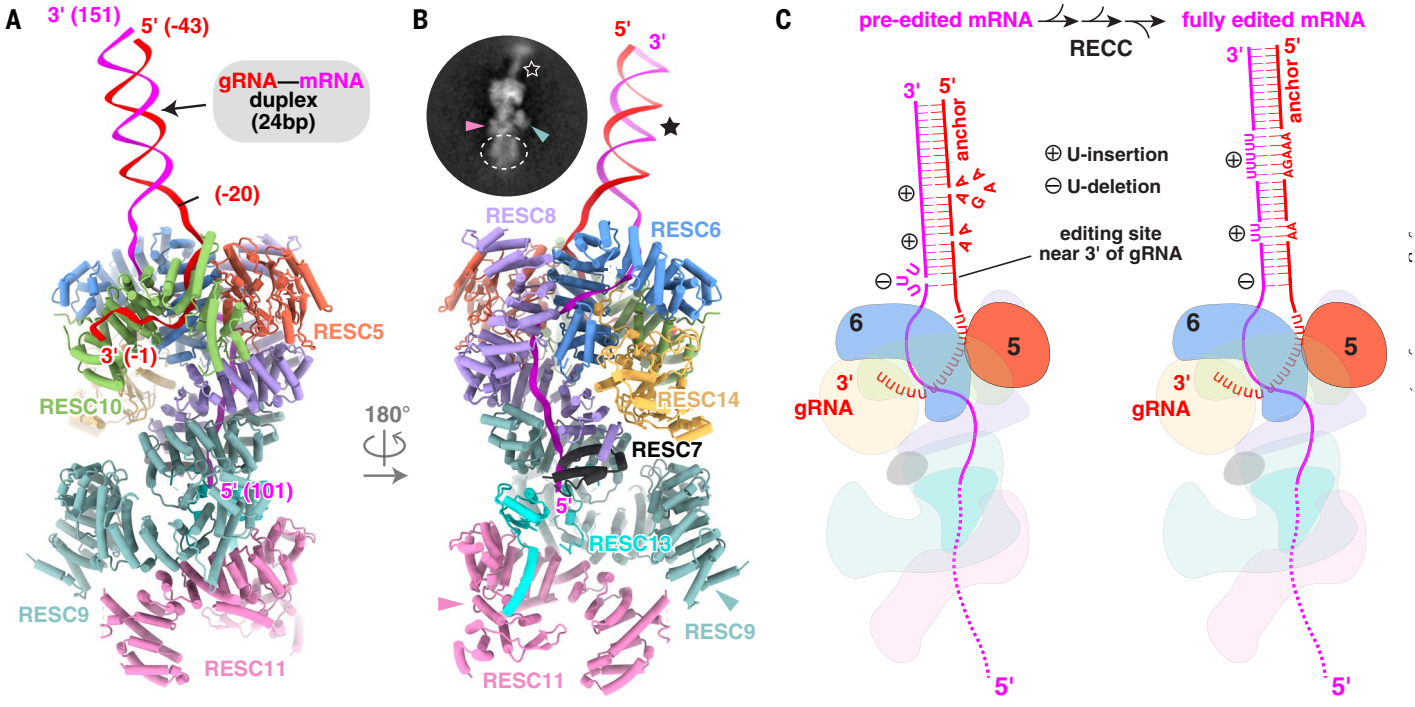



**Fig. 5. Overall architecture of gRNA-mRNA hybrid exiting RESC-B.** (**A** and **B**) Flipped views of the 24-bp duplex between gRNA and edited mRNA protruding out of RESC5, RESC6, RESC8, and RESC10. Three-dimensional (3D) projection of the cryo-EM map in (B) shows the gRNA-mRNA duplex (black star) and putative RESC12/12A (dashed circle). This 2D class average shares the same view as the RESC-B structure shown underneath. (**C**) A model of RNA-editing substrate. Mismatches in the duplex between gRNA and pre-edited mRNA delineate U-insertion (+) and U-deletion (−) editing sites (left). Completion of editing events extends the double-stranded region initially formed by the anchor through the guiding gRNA parts until full complementarity is achieved (right).

RESC-B complex tethers gRNA's 3′ end and pre-edited transcript to facilitate mRNA recognition by the anchor. The resultant imperfect mRNA-gRNA duplex protrudes from RESC-B surface (Fig. 5), exposing editing sites defined by a single gRNA to transiently binding and dissociating insertion- and deletion-specific RECC complexes (*48*, *49*) and, possibly, to postediting transactions. For example, during pan-editing that involves multiple overlapping gRNAs, the initiating gRNA directs sequence changes that create a binding site for the next one (*6*). However, initiating gRNA must be dislodged from the edited mRNA before the next one anneals, and this may involve RESC-A and RESC-B cycling or RESC-B reduction into RESC-C. It is also plausible that translation initiation would require gRNA clearing from the initiation codon created by editing (*50*, *51*).

Because RNA processing complexes typically undergo transient rearrangements implemented by molecular motors, the structures of stable particles suggest that RESC-A remodeling into RESC-B extracts gRNA's 5′ end from the RESC2 triphosphate tunnel and ejects subunits RESC1-4 to unfold the gRNA hairpin. To transform RESC-A to RESC-B, the following displacements ought to occur on the basis of the structural comparison by shared components RESC5 and RESC6 (Fig. 6A, left): RESC10 and RESC14 will roughly occupy the position of RESC2 and RESC1, respectively (Fig. 6A, middle); RESC7 and RESC8 will replace RESC3 and RESC4 (Fig. 6A, right); and RESC2 displacement yields the superhelical inner surface of RESC6, which binds to mRNA's 5′ segment in RESC-B (Fig. 6B).

Integration of biochemical data and structural analyses enables understanding of RESC assembly and rearrangements during the editing process (Fig. 6C). Separation of RESC2-associated complexes detects heterogeneous species with apparent molecular masses from ~280 kDa to ~1.2 MDa (Fig. 1) and demonstrates that the most abundant ~280 kDa particle lacks gRNAs. This correlates with the RESC-A cryo-EM structure, indicating that gRNA is sequestered into RESC-A during assembly from RESC1/2, RESC5/6, and RESC3 and RESC4 (Fig. 6C, left middle). Our results demonstrate a prominent RNA-dependent interaction between RESC-A and KREH2 RNA helicase, implicating this molecular motor in

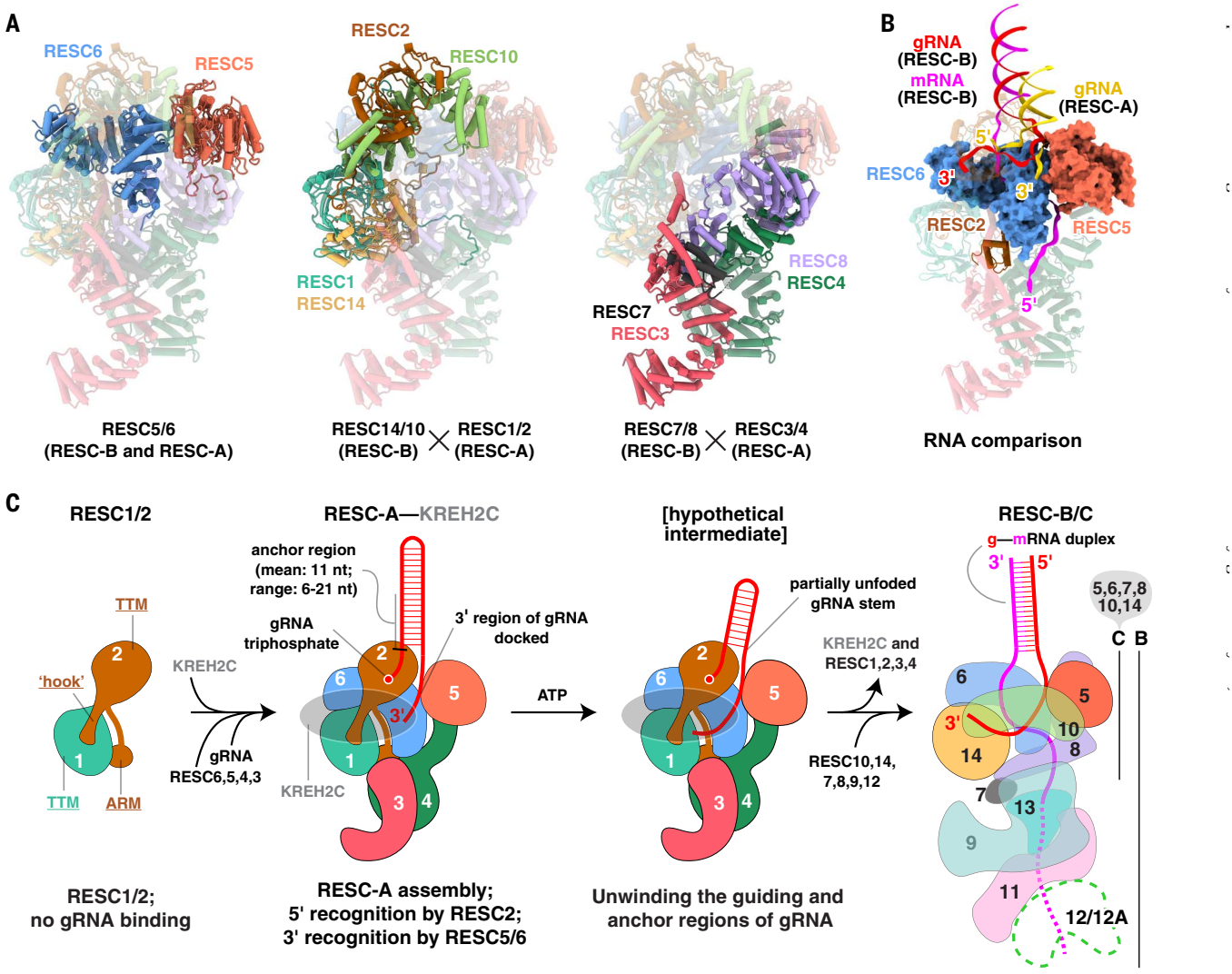



**Fig. 6. Model of transition from RESC-A to RESC-B.** (**A**) Potential clashes resolved by remodeling of RESC-A to RESC-B. RESC5 and RESC6 are well aligned in both complexes (left); RESC14/10 in RESC-B replace RESC1/2 in RESC-A (middle), and RESC7/8 in RESC-B replace RESC3/4 in RESC-A (right). RESC9 and RESC12 positioned in RESC-B midsection do not collide with RESC-A proteins. (**B**) Comparison of RESC-A and RESC-B highlights the preservation of RESC5/6 (surface) and repositioning of gRNA's single-stranded 3′ end (ribbon). RESC-B proteins are not shown for better visualization. (**C**) Proposed model of gRNA-stabilizing RESC-A assembly and KREH2C-induced remodeling into the editing-competent RESC-B.

RESC-A assembly (Fig. 1, A to F). Conversely, undetectable KREH2 levels in RESC9 and RESC14 complexes suggest that the helicase dissociates from RESC-A during remodeling into RESC-B (table S1). In RESC-A, gRNA anchor and guiding regions fold into a hairpin, whereas the uridylated 3′ end is affixed in the crevice between RESC5 and RESC6 (Fig. 6C, left middle). A single-stranded 3′ end is required for recognition by KREH2, which is capable of unwinding double-stranded RNA (*52*, *53*) and may unfold the hairpin to initiate RESC-A remodeling into RESC-B (Fig. 6C, right middle). In this scenario, recognition by KREH2 may select RESC5/6-bound gRNAs with longer U tails (Fig. 3J) as a quality checkpoint of RESC-A assembly. Collectively, our findings highlight KREH2 helicase as the probable remodeler facilitating the transition from RESC-A to RESC-B. We postulate that RESC-B is the editing-competent substrate of the RECC catalytic complex, for which our structures reveal an unobstructed approach path to the gRNA-mRNA hybrid.

## Materials and methods summary

### *Purification of the T. brucei* RESC complex

For protein affinity purification and enhanced in vivo UV cross-linking/affinity purification/RNA sequencing (eCLAP), we introduced a CTS tag (combinatorial affinity tag consisting of 10 histidines, protein C peptide epitope, and Twin-Strep peptide) (fig. S1A) at the C-termini of established RESC proteins (*8*). Protein-RNA adducts were purified by means of sequential streptavidin and metal affinity pulldowns and SDS-PAGE. RNA fragments were released through protease digestion and sequenced, as described in supplementary materials. Protein complexes were purified with tandem affinity chromatography from enriched mitochondrial fraction of procyclic (insect) form of *T. brucei* Lister 427 strain and analyzed as described in the supplementary materials.

### *EM grid preparation, imaging, data processing, and modeling*

Cryo-EM grids of both RNase-treated and mock-treated samples were plunge-frozen into precooled, liquefied ethane using Vitrobot Mark IV, screened with FEI TF20, and imaged by using an FEI Titan Krios electron microscope. Data processing of the RNase-treated dataset yielded three final maps with average resolution (FSC 0.143) of 3.4, 3.7, and 3.3 Å, respectively. These maps were identified through the CryoID approach and modeled as RESC1-6–containing RESC-A (3.7 Å), RESC5-14–containing RESC-B (3.4 Å), and RESC5-8,10,14–containing RESC-C (3.3 Å). In combination with the mock-treated RESC-A/B/C, we were able to identify a "closed" gRNA in RESC-A and gRNA-mRNA duplex in RESC-B/C. Details of EM grid preparation, imaging, data processing, and modeling are described in the supplementary materials.

## REFERENCES AND NOTES


1. R. Benne *et al*., Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**, 819–826 (1986). doi: 10.1016/0092-8674(86)90063-2; pmid: 3019552
2. J. E. Feagin, J. M. Abraham, K. Stuart, Extensive editing of the cytochrome c oxidase III transcript in Trypanosoma brucei. *Cell* **53**, 413–422 (1988). doi: 10.1016/0092-8674(88)90161-4; pmid: 2452697
3. B. Blum, N. Bakalara, L. Simpson, A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**, 189–198 (1990). doi: 10.1016/0092-8674(90)90735-W; pmid: 1688737
4. T. Suematsu *et al*., Antisense Transcripts Delimit Exonucleolytic Activity of the Mitochondrial 3′ Processome to Generate Guide RNAs. *Mol. Cell* **61**, 364–378 (2016). doi: 10.1016/j.molcel.2016.01.004; pmid: 26833087
5. B. Blum, L. Simpson, Guide RNAs in kinetoplastid mitochondria have a nonencoded 3′ oligo(U) tail involved in recognition of the preedited region. *Cell* **62**, 391–397 (1990). doi: 10.1016/0092-8674(90)90375-O; pmid: 1695552
6. D. A. Maslov, L. Simpson, The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* **70**, 459–467 (1992). doi: 10.1016/0092-8674(92)90170-H; pmid: 1379519
7. I. Aphasizheva *et al*., RNA binding and core complexes constitute the U-insertion/deletion editosome. *Mol. Cell. Biol.* **34**, 4329–4342 (2014). doi: 10.1128/MCB.01075-14; pmid: 25225332
8. I. Aphasizheva *et al*., Lexis and Grammar of Mitochondrial RNA Processing in Trypanosomes. *Trends Parasitol.* **36**, 337–355 (2020). doi: 10.1016/j.pt.2020.01.006; pmid: 32191849
9. S. M. McDermott, J. Luo, J. Carnes, J. A. Ranish, K. Stuart, The Architecture of Trypanosoma brucei editosomes. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6476–E6485 (2016). doi: 10.1073/pnas.1610177113; pmid: 27708162
10. J. Weng *et al*., Guide RNA-binding complex from mitochondria of trypanosomatids. *Mol. Cell* **32**, 198–209 (2008). doi: 10.1016/j.molcel.2008.08.023; pmid: 18951088
11. H. Hashimi, Z. Cicová, L. Novotná, Y. Z. Wen, J. Lukes, Kinetoplastid guide RNA biogenesis is dependent on subunits of the mitochondrial RNA binding complex 1 and mitochondrial RNA polymerase. *RNA* **15**, 588–599 (2009). doi: 10.1261/rna.1411809; pmid: 19228586
12. A. K. Panigrahi *et al*., Mitochondrial complexes in Trypanosoma brucei: A novel complex and a unique oxidoreductase complex. *Mol. Cell. Proteomics* **7**, 534–545 (2008). doi: 10.1074/mcp.M700430-MCP200; pmid: 18073385
13. M. L. Ammerman *et al*., Architecture of the trypanosome RNA editing accessory complex, MRB1. *Nucleic Acids Res.* **40**, 5637–5650 (2012). doi: 10.1093/nar/gks211; pmid: 22396527
14. L. K. Read, J. Lukeš, H. Hashimi, Trypanosome RNA editing: The complexity of getting U in and taking U out. *Wiley Interdiscip. Rev. RNA* **7**, 33–51 (2016). doi: 10.1002/wrna.1313; pmid: 26522170
15. L. N. Rusché, J. Cruz-Reyes, K. J. Piller, B. Sollner-Webb, Purification of a functional enzymatic editing complex from Trypanosoma brucei mitochondria. *EMBO J.* **16**, 4069–4081 (1997). doi: 10.1093/emboj/16.13.4069; pmid: 9233816
16. A. K. Panigrahi *et al*., Identification of novel components of Trypanosoma brucei editosomes. *RNA* **9**, 484–492 (2003). doi: 10.1261/rna.2194603; pmid: 12649499
17. R. Aphasizhev *et al*., Isolation of a U-insertion/deletion editing complex from Leishmania tarentolae mitochondria. *EMBO J.* **22**, 913–924 (2003). doi: 10.1093/emboj/cdg083; pmid: 12574127
18. J. Deng, N. L. Ernst, S. Turley, K. D. Stuart, W. G. Hol, Structural basis for UTP specificity of RNA editing TUTases from Trypanosoma brucei. *EMBO J.* **24**, 4007–4017 (2005). doi: 10.1038/sj.emboj.7600861; pmid: 16281058
19. J. Deng, A. Schnaufer, R. Salavati, K. D. Stuart, W. G. Hol, High resolution crystal structure of a key editosome enzyme from Trypanosoma brucei: RNA editing ligase 1. *J. Mol. Biol.* **343**, 601–613 (2004). doi: 10.1016/j.jmb.2004.08.041; pmid: 15465048
20. J. Stagno, I. Aphasizheva, J. Bruystens, H. Luecke, R. Aphasizhev, Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. *J. Mol. Biol.* **399**, 464–475 (2010). doi: 10.1016/j.jmb.2010.04.021; pmid: 20403364
21. M. Wu *et al*., Structures of a key interaction protein from the Trypanosoma brucei editosome in complex with single domain antibodies. *J. Struct. Biol.* (2010). pmid: 20969962
22. Y. J. Park *et al*., The structure of the C-terminal domain of the largest editosome interaction protein and its role in promoting RNA binding by RNA-editing ligase L2. *Nucleic Acids Res.* **40**, 6966–6977 (2012). doi: 10.1093/nar/gks369; pmid: 22561373
23. P. L. Shaw *et al*., Structures of the T. brucei kRNA editing factor MRB1590 reveal unique RNA-binding pore motif contained within an ABC-ATPase fold. *Nucleic Acids Res.* **43**, 7096–7109 (2015). doi: 10.1093/nar/gkv647; pmid: 26117548
24. M. M. Golas *et al*., Snapshots of the RNA editing machine in trypanosomes captured at different assembly stages in vivo. *EMBO J.* **28**, 766–778 (2009). doi: 10.1038/emboj.2009.19; pmid: 19197238
25. F. Li *et al*., Structure of the core editing complex (L-complex) involved in uridine insertion/deletion RNA editing in trypanosomatid mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 12306–12310 (2009). doi: 10.1073/pnas.0901754106; pmid: 19590014
26. C. M. Ho *et al*., Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nat. Methods* **17**, 79–85 (2020). doi: 10.1038/s41592-019-0637-y; pmid: 31768063
27. L. G. Dolce, Y. Nesterenko, L. Walther, F. Weis, E. Kowalinski, Structural basis for guide RNA selection by the RESC1-RESC2 complex. *Nucleic Acids Res.* **51**, 4602–4612 (2023). doi: 10.1093/nar/gkad217; pmid: 36999600
28. S. Dixit *et al*., Differential Binding of Mitochondrial Transcripts by MRB8170 and MRB4160 Regulates Distinct Editing Fates of Mitochondrial mRNA in Trypanosomes. *mBio* **8**, e02288-16 (2017). doi: 10.1128/mBio.02288-16; pmid: 28143982
29. L. Kafková *et al*., Functional characterization of two paralogs that are novel RNA binding proteins influencing mitochondrial transcripts of Trypanosoma brucei. *RNA* **18**, 1846–1861 (2012). doi: 10.1261/rna.033852.112; pmid: 22898985
30. M. A. Andrade, C. Petosa, S. I. O'Donoghue, C. W. Müller, P. Bork, Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* **309**, 1–18 (2001). doi: 10.1006/jmbi.2001.4624; pmid: 11491282
31. A. F. Neuwald, T. Hirano, HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. *Genome Res.* **10**, 1445–1452 (2000). doi: 10.1101/gr.147400; pmid: 11042144
32. C. D. Lima, L. K. Wang, S. Shuman, Structure and mechanism of yeast RNA triphosphatase: An essential component of the mRNA capping apparatus. *Cell* **99**, 533–543 (1999). doi: 10.1016/S0092-8674(00)81541-X; pmid: 10589681
33. S. S. Leung, D. J. Koslowsky, Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res.* **27**, 778–787 (1999). doi: 10.1093/nar/27.3.778; pmid: 9889273
34. L. Reifur, D. J. Koslowsky, Trypanosoma brucei ATPase subunit 6 mRNA bound to gA6-14 forms a conserved three-helical structure. *RNA* **14**, 2195–2211 (2008). doi: 10.1261/rna.1144508; pmid: 18772247
35. S. Cooper, E. S. Wadsworth, A. Schnaufer, N. J. Savill, Organization of minicircle cassettes and guide RNA genes in *Trypanosoma brucei*. *RNA* **28**, 972–992 (2022). doi: 10.1261/rna.079022.121; pmid: 35414587
36. J. L. Mattiacio, L. K. Read, Roles for TbDSS-1 in RNA surveillance and decay of maturation by-products from the 12S rRNA locus. *Nucleic Acids Res.* **36**, 319–329 (2008). doi: 10.1093/nar/gkm690; pmid: 18032430
37. S. L. Zimmer, S. M. McEvoy, J. Li, J. Qu, L. K. Read, A novel member of the RNase D exoribonuclease family functions in mitochondrial guide RNA metabolism in Trypanosoma brucei. *J. Biol. Chem.* **286**, 10329–10340 (2011). doi: 10.1074/jbc.M110.152439; pmid: 21252235
38. F. M. Sement *et al*., Transcription initiation defines kinetoplast RNA boundaries. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10323–E10332 (2018). doi: 10.1073/pnas.1808981115; pmid: 30333188
39. T. R. Cech, RNA editing: World's smallest introns? *Cell* **64**, 667–669 (1991). doi: 10.1016/0092-8674(91)90494-J; pmid: 1997201
40. R. Aphasizhev, I. Aphasizheva, L. Simpson, A tale of two TUTases. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10617–10622 (2003). doi: 10.1073/pnas.1833120100; pmid: 12954983
41. B. M. Foda, K. M. Downey, J. C. Fisk, L. K. Read, Multifunctional G-rich and RRM-containing domains of TbRGG2 perform

separate yet essential functions in trypanosome RNA editing. *Eukaryot. Cell* **11**, 1119–1131 (2012). doi: 10.1128/EC.00175-12; pmid: 22798390
42. E. L. Van Nostrand *et al*., Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP). *Methods Mol. Biol.* **1648**, 177–200 (2017). doi: 10.1007/978-1-4939-7204-3_14; pmid: 28766298
43. N. M. McAdams, R. M. Simpson, R. Chen, Y. Sun, L. K. Read, MRB7260 is essential for productive protein-RNA interactions within the RNA editing substrate binding complex during trypanosome RNA editing. *RNA* **24**, 540–556 (2018). doi: 10.1261/rna.065169.117; pmid: 29330168
44. N. M. McAdams *et al*., MRB10130 is a RESC assembly factor that promotes kinetoplastid RNA editing initiation and progression. *RNA* **25**, 1177–1191 (2019). doi: 10.1261/rna.071902.119; pmid: 31221726
45. R. P. Igo Jr., S. D. Lawson, K. Stuart, RNA sequence and base pairing effects on insertion editing in Trypanosoma brucei. *Mol. Cell. Biol.* **22**, 1567–1576 (2002). doi: 10.1128/MCB.22.5.1567-1576.2002; pmid: 11839822
46. S. D. Lawson, R. P. Igo Jr., R. Salavati, K. D. Stuart, The specificity of nucleotide removal during RNA editing in *Trypanosoma brucei*. *RNA* **7**, 1793–1802 (2001). doi: 10.1017/S135583820101055X; pmid: 11780635
47. R. P. Igo Jr., S. S. Palazzo, M. L. Burgess, A. K. Panigrahi, K. Stuart, Uridylate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol. Cell. Biol.* **20**, 8447–8457 (2000). doi: 10.1128/MCB.20.22.8447-8457.2000; pmid: 11046141
48. J. R. Trotter, N. L. Ernst, J. Carnes, B. Panicucci, K. Stuart, A deletion site editing endonuclease in Trypanosoma brucei. *Mol. Cell* **20**, 403–412 (2005). doi: 10.1016/j.molcel.2005.09.016; pmid: 16285922
49. J. Carnes, J. R. Trotter, N. L. Ernst, A. Steinberg, K. Stuart, An essential RNase III insertion editing endonuclease in Trypanosoma brucei. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16614–16619 (2005). doi: 10.1073/pnas.0506133102; pmid: 16269544
50. T. Lenarčič *et al*., Mitoribosomal small subunit maturation involves formation of initiation-like complexes. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2114710118 (2022). doi: 10.1073/pnas.2114710118; pmid: 35042777
51. J. E. Feagin, J. M. Shaw, L. Simpson, K. Stuart, Creation of AUG initiation codons by addition of uridines within cytochrome b transcripts of kinetoplastids. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 539–543 (1988). doi: 10.1073/pnas.85.2.539; pmid: 2448777
52. V. Kumar *et al*., REH2C Helicase and GRBC Subcomplexes May Base Pair through mRNA and Small Guide RNA in Kinetoplastid Editosomes. *J. Biol. Chem.* **291**, 5753–5764 (2016). doi: 10.1074/jbc.M115.708164; pmid: 26769962
53. A. Hernandez *et al*., REH2 RNA helicase in kinetoplastid mitochondria: Ribonucleoprotein complexes and essential motifs for unwinding and guide RNA (gRNA) binding. *J. Biol. Chem.* **285**, 1220–1228 (2010). doi: 10.1074/jbc.M109.051862; pmid: 19850921
54. L. Zhang, bioliyezhang/RESC_paper, version v1.0.0, Zenodo (2023); doi: 10.5281/zenodo.7930544

## ACKNOWLEDGMENTS

We thank Y. He and Y.-T. Liu for evaluating samples at the early stage of this project; W. Lan for initial atomic model–building efforts for part of RESC8, RESC9, and RESC14 in RESC-B structure obtained from RNase-treated isolate; A. Vacas for assistance with transgenic cell lines; and S. Shuman and J. C. Samuelson for discussions. We also thank the high-performance computing (HPC) platform of ShanghaiTech University for support. **Funding:** NIH grants R35GM145249 and R01GM074830 were given to L.H., RO1AI113157 to I.A., RO1AI101057 and R01AI152408 to R.A., and R01GM071940 to Z.H.Z. L.Z. is supported by National Natural Science Foundation of China (grant 31871332). We acknowledge the use of resources in the Electron Imaging Center for Nanomachines, supported by UCLA, and grants from NIH (S10RR23057, S10OD018111, and U24GM116792) and NSF (DBI-1338135 and DMR-1548924). **Author contributions:** Project initiation and supervision: Z.H.Z. and R.A.; Transgenic cell lines generation: H.W. and Q.L.; Biochemical analyses: H.W., I.A., and T.S.; Mass spectrometry: C.Y. and L.H.; Cryo-EM sample preparation: H.W.; EM sample evaluation and optimization: A.L.C., X.L., and S.L.; Cryo-EM imaging for RNase-treated samples: A.L.C. and X.L.; Data processing for RNase-treated samples: S.L., X.L. and A.L.C.; Cryo-EM imaging and data processing for mock-treated samples: S.L.; CryoID: X.L. and Z.L.; Protein atomic modeling: Z.L., J.K.J.L., H.W., J.J.H., and S.L.; Nucleotide identification and modeling: S.L.; eCLAP data generation: I.A.; eCLAP data analysis: F.Z., X.Z., and L.Z.; Data interpretation: S.L., Z.H.Z., R.A., I.A., J.J.H., J.K.J.L., and L.Z.; Figure preparation: S.L., J.K.J.L., J.J.H., I.A., F.Z., X.Z., L.Z., R.A., and Z.H.Z.; Movie preparation: J.J.H. and J.K.J.L.; Manuscript writing: S.L., L.Z., I.A., Z.H.Z., and R.A. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The cryo-EM maps and atomic coordinates have been deposited to the EMDB (EMD-29308, EMD-29314, EMD-29305, EMD-29311, EMD-29316, EMD-29306 for RESC-A, -B, and -C with and without RNase treatment) and PDB (8FNC, 8FNI, 8FN4, 8FNF, 8FNK, 8FN6 for RESC-A, -B, and -C, with and without RNase treatment) databases, respectively, and will be released at time of publication. The raw FASTQ files produced by eCLAP experiments are deposited into the NCBI SRA database under accession number PRJNA879925. Scripts for data analysis and visualization are available at Zenodo (*54*). Mass spectrometry data have been deposited to the ProteomeXchange Consortium with the dataset identifier PXD039156. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg4725
Materials and Methods
Figs. S1 to S16
Tables S1 to S6
References (*55*–*77*)
MDAR Reproducibility Checklist
Movies S1 to S3

View/request a protocol for this paper from *Bio-protocol*.

Submitted 28 December 2022; accepted 17 May 2023
10.1126/science.adg4725

RESEARCH ARTICLE SUMMARY

STRUCTURAL BIOLOGY

# Principles of human pre-60*S* biogenesis

**Arnaud Vanden Broeck* and Sebastian Klinge***

**INTRODUCTION:** Ribosomes, two-subunit RNA-protein nanomachines, translate messenger RNA (mRNA) into proteins in all living organisms. The small ribosomal subunit (40*S*) is responsible for decoding of mRNA, whereas the large subunit (60*S*) catalyzes peptide bond formation. Assembly of both ribosomal subunits in human cells is initiated in the nucleolus, followed by nuclear and cytoplasmic maturation, and requires more than 200 ribosome assembly factors catalyzing modification, processing, and folding of ribosomal RNA (rRNA). During nucleolar assembly, the large ribosomal subunit precursor (pre-60*S*) is assembled from a 5*S* rRNA and a 32*S* pre-rRNA precursor, the latter containing the 5.8*S* rRNA and the 28*S* rRNA joined through an internal transcribed spacer 2 (ITS2). Because of an inability to isolate early endogenous pre-60*S* assembly intermediates, insights into human large ribosomal subunit assembly are so far limited to very late nuclear states. The mechanism underlying nucleolar assembly and nuclear maturation of human pre-60*S* particles thus remains unknown.

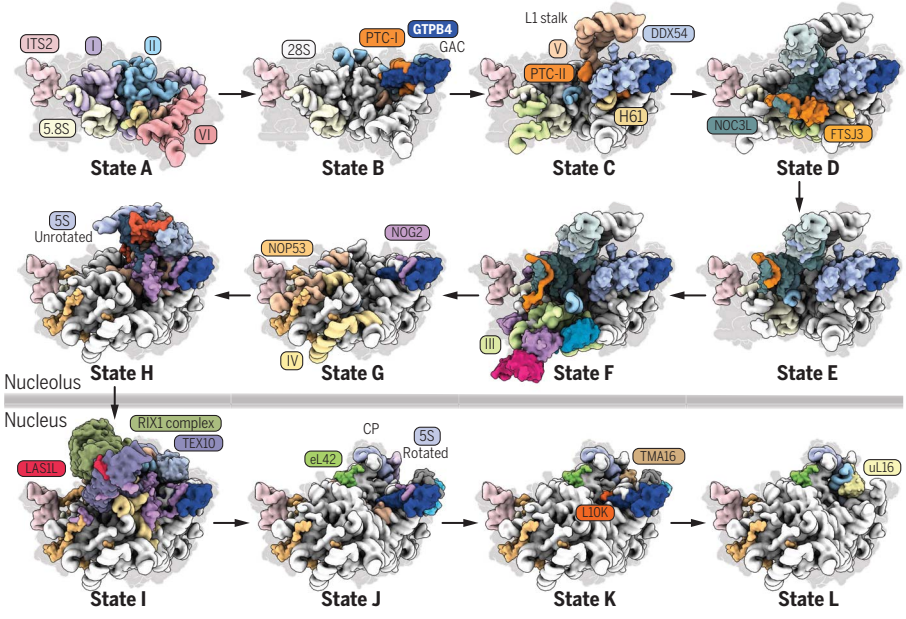
**RATIONALE:** We set out to elucidate the early stages of human pre-60*S* assembly at a mechanistic level. To this end, we have combined human genome editing and biochemistry to permeabilize human nucleoli and nuclei to isolate intact endogenous pre-60*S* assembly intermediates for subsequent structural characterization by cryo–electron microscopy (cryo-EM). To functionally study rRNA processing, we developed an in vivo recombinant ribosome assembly assay using an engineered human rDNA locus to investigate whether rRNA elements within ITS2 and the 28*S* rRNA are required for large ribosomal subunit biogenesis.

**RESULTS:** In this study, we biallelically affinity tagged the ITS2-associated assembly factor MK67I and determined 24 cryo-EM structures of human pre-60*S* assembly intermediates at resolutions of 2.5 to 3.2 Å. Within this structural landscape, we observed several parallel assembly pathways and identified eight main nucleolar assembly states and four main nuclear maturation states. In the nucleolus, the structures highlight how protein interaction hubs tether assembly factor complexes to the maturing pre-60*S* particles and how GTPases and ATPases, such as DEAD-box RNA helicases, can couple irreversible nucleotide hydrolysis steps to the installation of functional centers. After exiting the nucleolus, the maturing pre-60*S* particles demonstrate how the structural plasticity of the pre-rRNA is interrogated by an ensemble of assembly factors including the rixosome to couple rRNA conformational changes with RNA exosome–mediated degradation of ITS2. Later nuclear pre-60*S* states reveal how assembly factors interchange is used to drive irreversible fine tuning of functional centers. Additionally, by using engineered human rRNAs, we have identified elements of ITS2 and the 28*S* rRNA that are critical to generating mature large ribosomal subunits in human cells. The high resolution of our cryo-EM reconstructions further allowed us to visualize many previously mapped chemical modifications of the human 28*S* rRNA, providing a foundation with which to study both universally conserved chemical modifications as well as species-specific adaptations as a function of ribosome assembly.

**CONCLUSION:** Our ensemble of cryo-EM structures, together with our engineered ribosome assembly assay, now provide a high-resolution perspective on nucleolar assembly and nuclear maturation of the human large ribosomal subunit. Our study rationalizes decades of genetic and biochemical data and reveal new principles of eukaryotic ribosome assembly. ■

Laboratory of Protein and Nucleic Acid Chemistry, The Rockefeller University, New York, NY 10065, USA.
*Corresponding author. Email: klinge@rockefeller.edu (S.K.); avanden@rockefeller.edu (A.V.B.)
Cite this article as A. Vanden Broeck, S. Klinge, *Science* **381**, eadh3892 (2023). DOI: 10.1126/science.adh3892

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adh3892




**Principles of human pre-60*S* biogenesis.** Human pre-60*S* assembly pathway with crown view of states A to L. During nucleolar assembly, states A to C uncover the formation of the peptidyl transferase center (PTC) in two stages (PTC-I and PTC-II). Toward nuclear maturation, states F to H highlight major remodeling events to accommodate the immature central protuberance (CP) containing the unrotated 5*S* rRNA. The transition from state H to I reveals the coupling of 5*S* rotation to ITS2 processing. Late nuclear maturation steps illustrated by states J to L include the formation of the E site and fine tuning of the CP and PTC. Newly folded rRNA elements are color-coded. Key assembly factors or ribosomal proteins are shown.

RESEARCH ARTICLE

STRUCTURAL BIOLOGY

# Principles of human pre-60*S* biogenesis

**Arnaud Vanden Broeck* and Sebastian Klinge***

**During the early stages of human large ribosomal subunit (60*S*) biogenesis, an ensemble of assembly factors establishes and fine-tunes the essential RNA functional centers of pre-60*S* particles by an unknown mechanism. Here, we report a series of cryo–electron microscopy structures of human nucleolar and nuclear pre-60*S* assembly intermediates at resolutions of 2.5 to 3.2 angstroms. These structures show how protein interaction hubs tether assembly factor complexes to nucleolar particles and how guanosine triphosphatases and adenosine triphosphatase couple irreversible nucleotide hydrolysis steps to the installation of functional centers. Nuclear stages highlight how a conserved RNA-processing complex, the rixosome, couples large-scale RNA conformational changes with pre–ribosomal RNA processing by the RNA degradation machinery. Our ensemble of human pre-60*S* particles provides a rich foundation with which to elucidate the molecular principles of ribosome formation.**

Protein synthesis in all cells requires ribosomes, two-subunit RNA-protein molecular machines in which the RNA component is responsible for both decoding of messenger RNA within the ribosomal small subunit (40*S*) as well as peptide bond formation within the ribosomal large subunit (60*S*). Since the discovery of ribosome assembly in human cells, >200 ribosome assembly factors have been identified that catalyze the modification, processing, and folding of ribosomal RNA (rRNA), initially in the nucleolus and subsequently in the nucleus and cytoplasm (*1*–*3*). The large ribosomal subunit rRNAs (5*S*, 5.8*S*, and 28*S* rRNA) are generated from two transcripts: a transcript containing the 5*S* rRNA and a separate 47*S* pre-rRNA that contains the 5.8*S* and 28*S* rRNA in addition to the rRNA of the small subunit. After the transcription of the 47*S* pre-rRNA in the nucleolus and a cleavage reaction separating precursors of both subunits, the first stable nucleolar large subunit pre-rRNA is a 32*S* precursor in which the 5.8*S* and 28*S* rRNAs are joined by internal transcribed spacer 2 (ITS2).

Traditionally, eukaryotic 60*S* assembly has been studied in the yeast system [see (*4*) and references therein and (*5*–*10*)], which has offered several molecular snapshots of early nucleolar assembly intermediates from *Saccharomyces cerevisiae* (*11*–*13*). However, the mechanisms that drive early assembly forward remain unknown. The absence of key transition states has so far precluded a mechanistic understanding of how essential energy-consuming enzymes such as adenosine triphosphatases (ATPases), including DEAD/H-box helicases and AAA+-ATPases, as well as guanosine triphosphatases (GTPases) can bring about irreversible transitions and the formation of catalytic centers of the eukaryotic large ribosomal subunit, such as the peptidyl transferase center (PTC) (*14*, *15*). Thus, the mechanisms that enable early eukaryotic large ribosomal subunit assembly to proceed in a controlled and unidirectional manner remain poorly understood. Toward the end of nucleolar and beginning of nuclear maturation, pre-60*S* assembly intermediates already contain functional centers, including the PTC, the GTPase-associated center (GAC), the L1 stalk, as well as a central protuberance (CP) containing an unrotated 5*S* ribonucleoprotein (RNP; Nog2 particle state 1) (*16*). However, these functional centers have not yet reached their mature conformations, requiring an ensemble of nuclear ribosome assembly factors for their completion. Although several structures covering nuclear stages of large ribosomal subunit assembly in yeast have been described (*5*–*7*, *16*, *17*), major gaps in our mechanistic understanding of this process remain. An essential transition during nuclear pre-60*S* maturation is the rotation of the 5*S* RNP during conversion of a Nog2-bound particle (*16*) to a Rix1-Midasin bound particle (*7*), which is catalyzed by Midasin (Mdn1). Only after the rotation of the 5*S* RNP does ITS2 processing result in the mature 5.8*S* rRNA, but how these two irreversible events are linked has remained unclear. Consistent with functional coupling between 5*S* RNP rotation and ITS2 processing, mutations or depletions of assembly factors involved in 5*S* RNP rotation or ITS2 processing result in the accumulation of either unprocessed pre-rRNA or 5.8*S* precursors containing parts of ITS2 (*18*–*22*). The accumulation of the same pre-RNA species is also observed upon depletion of the endonuclease cleaving ITS2 (yeast Las1, human LAS1L) and components of the RNA exosome, respectively (*23*, *24*). Although these data suggest coupling of 5*S* RNP rotation with ITS2 cleavage and processing by Las1/LAS1L and the RNA exosome, the molecular basis of this mechanism remains elusive.

Insights into human large ribosomal subunit assembly are still in their infancy, because only a few structures of very late stages of pre-60*S* maturation have been obtained through the ectopic expression of an affinity-tagged nuclear export factor (*25*). Because of the biochemical properties of the human nucleolus, an inability to obtain endogenous early pre-60*S* assembly intermediates has so far limited our understanding of the molecular mechanisms responsible for their unidirectional assembly and maturation. Furthermore, within the human nuclear ribosome assembly pathway, it is unclear how an evolutionarily expanded multiprotein complex, the rixosome (*26*), can integrate different functional activities to ensure that cleavage of ITS2 by LAS1L (*27*) and the subsequent RNA exosome–mediated processing of ITS2 are efficiently coupled.

From an evolutionary perspective, the 32*S* pre-rRNA of the late nucleolar intermediates in humans is much larger than its yeast 27*S* counterpart because of the presence of several expansion segments in ITS2 and the 28*S* rRNA, and it is currently unknown if these expansions are required for ribosome assembly.

In this study, we have combined human genome editing and biochemistry to permeabilize human nucleoli and nuclei to determine high-resolution cryo–electron microscopy (cryo-EM) structures of 24 maturing human pre-60*S* particles at resolutions of 2.5 to 3.2 Å. Of these, eight main nucleolar structures elucidate how multiple levels of regulation are used to drive unidirectional 60*S* maturation by integrating tethered multiprotein complexes with ATPases, such as DEAD-box RNA helicases, and GTPases that catalyze key transitions during the assembly of functional centers. Four main nuclear structures reveal the mechanism by which 5*S* RNP rotation is coupled to ITS2 processing and how the human rixosome plays a central role in mediating these steps. By using engineered rRNAs in human cells, we have identified elements of the human ITS2 and 28*S* rRNA that are critical for ribosome assembly and nuclear export.

## Cryo-EM structures of maturing human pre-60*S* particles

To obtain structural insights into early stages of human large ribosomal subunit assembly, we have biallelically affinity tagged the endogenous ITS2-associated assembly factor MK67I (fig. S1, A to D). The biochemical affinity purification of MK67I enabled the direct isolation of human pre-60*S* particles containing ITS2 for cryo-EM analysis. These nucleolar and nuclear stages of assembly are associated with the expected nucleolar (32*S*) and nuclear (28*S*,

Laboratory of Protein and Nucleic Acid Chemistry, The Rockefeller University, New York, NY 10065, USA.
***Corresponding author. Email: klinge@rockefeller.edu (S.K.); avanden@rockefeller.edu (A.V.B.)**

12*S*, and 8*S*) pre-rRNAs (figs. S1, E to I, and S2). A total of 172,699 cryo-EM images were collected on a Titan Krios microscope equipped with a K3 detector. After initial data curation and extensive three-dimensional classification, we identified 12 nucleolar stages of assembly, of which the main eight states are here referred to as A to H, and 12 nuclear stages of assembly, of which the four main states are referred to as I to L (figs. S3 to S5). Focused classifications and refinements were used to obtain high-resolution composite maps (2.5 to 3.2 Å) (figs. S6 to S15 and tables S1 and S2). These reconstructions enabled the assignment and precise model building of human pre-rRNAs (5.8*S*, ITS2, 28*S*, and 5*S* rRNA), 41 ribosomal proteins, and 47 nucleolar or nuclear ribosome assembly factors (tables S3 and S4). Differences in the pre-rRNA folding states and the presence or absence of distinct assembly factors enabled the assignment of states A to H and states I to L into nucleolar and nuclear assembly pathways, respectively (Fig. 1, figs. S16 to S19, and movie 1).

We have identified parallel assembly pathways at the beginning of the nucleolar assembly (states A to D) characterized by the presence or absence of the MAK16-associated assembly factor complex [MAK16, WDR74 (yeast Nsa1), RPF1, RRP1 and the GTP-binding domain of GNL3 (yeast Nug1)] that are removed by NVL2 (*15*) (figs. S16 and S17). We also observed parallel maturation pathways with the stepwise

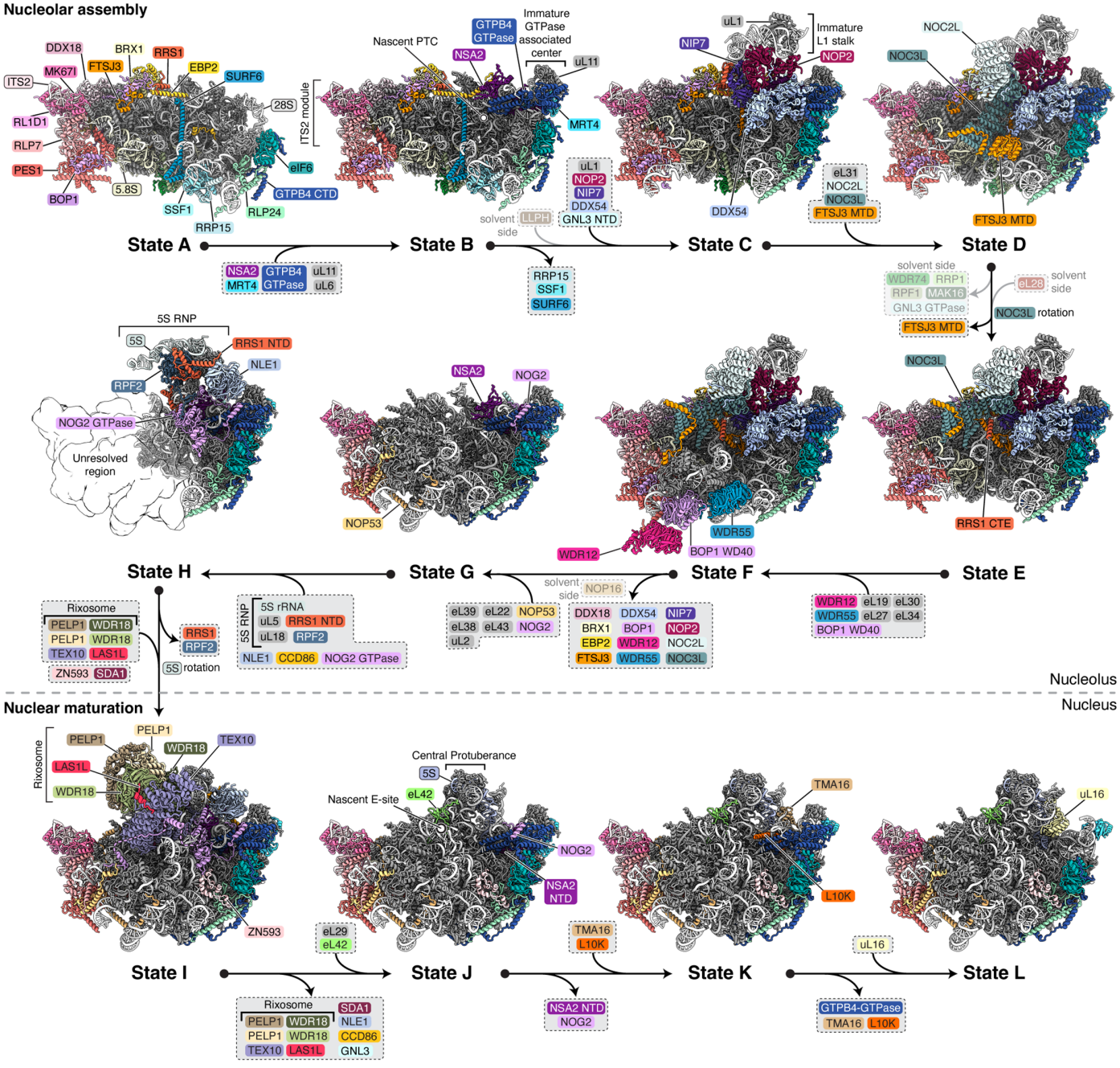



**Fig. 1. Nucleolar assembly and nuclear maturation of human pre-60*S* particles.** Simplified human pre-60*S* assembly pathway with crown view of atomic models of states A to L with ribosomal proteins in gray and color-coded pre-rRNAs and ribosome assembly factors. Proteins that became ordered or departed between different assembly states are indicated by arrows.

maturation of ITS2 elements from states I to L during nuclear stages of assembly (figs. S18 and S19). The ITS2-centered events can occur concurrently with other maturation events that focus on the fine-tuning of functional centers of progressing nuclear pre-60*S* particles. These include the cooperative formation of the E site (states J), the progressive fine-tuning of the PTC through irreversible exchange of assembly factors (states K), and the maturation of the P site with the installation of uL16 (states L).

The scale of our cryo-EM dataset, which contains ~15 million particles, of which ~2.4 million particles were chosen, allowed us to observe the progression of the human large ribosomal subunit assembly as a fine-grained series of molecular snapshots that enable a precise mechanistic explanation of how ribosome assembly factors act in concert, the identification of human-specific adaptations, and the visualization of covalent modifications of pre-rRNA as a function of ribosome assembly. This ensemble of states now provides a high-resolution perspective on nucleolar assembly and nuclear maturation of the human large ribosomal subunit as described below.

### High-level organization and hierarchy of pre-rRNA and protein incorporation

A unique principle of human pre-60*S* assembly is the early establishment of hubs that organize architectural pre-rRNA elements and tether ribosome assembly factor complexes to coordinate different steps of assembly (Fig. 2).

rRNA root helices, which form the origin of each of the six subdomains (I to VI) of the 28*S* rRNA, are already positioned in the first visualized posttranscriptional assembly state (state A) (fig. S20). The early establishment of root helices around the assembly factor complex SURF6-SSF1-RRP15 assists in the correct positioning of each rRNA domain, thereby facilitating subsequent folding events that occur either within a domain or across domains (Fig. 2A). Ribosome assembly factors are not only responsible for the correct placement of central architectural rRNA elements within state A, but also prevent premature folding of key rRNA elements (figs. S21 and S22).

Comparable to the organization of central elements of pre-rRNA, a protein-protein interaction hub is established in state A near domain I, which tethers ribosome assembly factor complexes to the maturing particles, thereby allowing for progression through subsequent stages of assembly (Fig. 2B). This protein-protein interaction hub is formed around an N-terminal region of BOP1 to which an extended BRX1-EBP2 complex and peptide elements of FTSJ3 and RRS1 are bound.

Previous studies have highlighted that eukaryotic ribosome assembly is controlled through both the TOR and casein kinase II pathways (*28*, *29*). We have identified two mammalian-specific posttranslational modifications in the N-terminal segment of BOP1 (phosphoserines 126 and 127), which are present within a sequence motif resembling casein kinase II substrates (*30*). The central location of these modified residues, which are present in all states containing BOP1 (states A to F), raises the possibility that this set of post-translational modifications enables modulation of human ribosome biogenesis within cellular metabolism through signaling pathways (fig. S22B).

During nucleolar assembly stages covering states A to F, the protein-protein interaction hub (Fig. 2B) performs three major organizing functions. First, the presence of a peptide of the RNA 2′-O-methyltransferase FTSJ3 (yeast Spb1) suggests that tethering enables 2′-O-methylation of 28*S* rRNA as well as recruitment of the NOC2L-NOC3L complex and ribosomal proteins at different stages of assembly (fig. S22B). By interacting with all of these proteins, tethered ribosome assembly factors and ribosomal proteins are already present before their involvement in subsequent stages of assembly. Second, the presence of a C-terminal extension of RRS1 provides a mechanism for recruitment and later incorporation of the L1 stalk and the 5*S* RNP through the RRS1-RPF2 complex that binds to 5*S* rRNA and ribosomal proteins uL18 (RPL5) and uL5 (RPL11) (Fig. 1 and fig. S22D). Third, EBP2 has dual functions that include both the prevention of premature rRNA folding as well as the stabilization of the NSA2-GTPB4 complex in state B (Figs. 1 and 2 and figs. S21B and S22C).

Collectively, the organization of pre-rRNA elements and ribosome assembly factor complexes in state A lays the foundation for many of the subsequent irreversible transitions that require the presence of already tethered assembly factors to efficiently stabilize new pre-rRNA folding states.

### GTPB4 catalyzes the initial installation of the peptidyl transferase center

Key transitions during ribosome assembly are controlled through irreversible steps that involve nucleotide hydrolysis, RNA cleavage, or the generation of force (*4*, *31*). An essential family of GTPases conserved in all domains of life that acts as a central regulator of large ribosomal subunit assembly includes the human protein GTPB4 (Nog1 in yeast, ObgE in *Escherichia coli*) (*16*, *32*–*34*). So far, the limiting resolution of endogenous assembly intermediates obtained from any species and the lack of transition states of maturing particles has precluded a precise mechanistic understanding of the GTPase cycle of GTPB4 and its homologs during assembly.

Whereas the C terminus of GTPB4 is already bound to eIF6 and RLP24 in state A (Fig. 3A), the guanosine diphosphate (GDP)–bound GTPB4 GTPase domain first appears in state B and

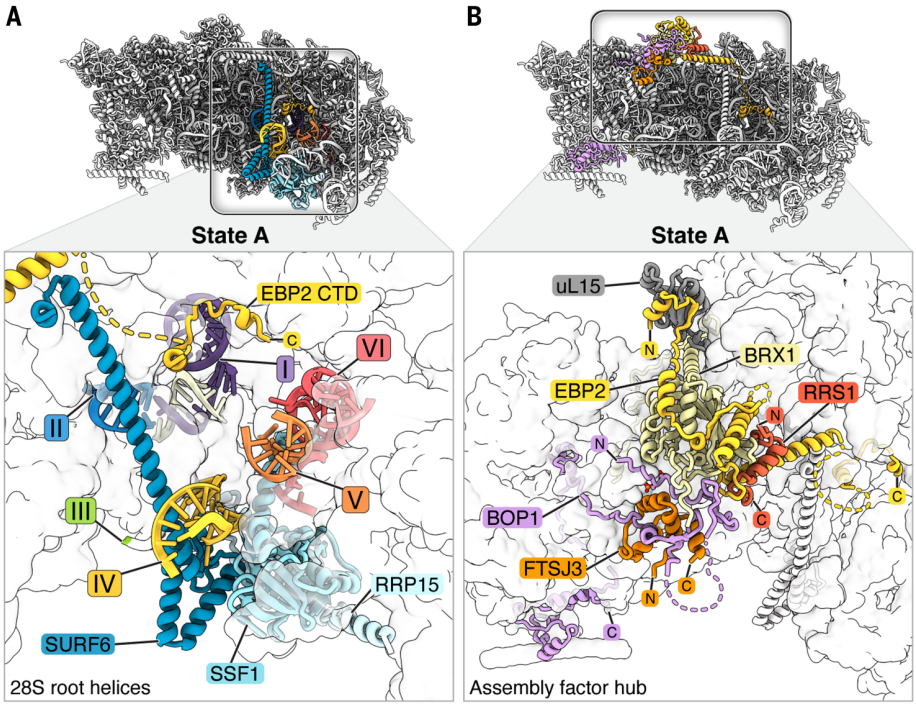



**Fig. 2. Early clustering of pre-rRNA and assembly factor modules.** (**A**) Crown view of state A showing color-coded root helices of domains I to VI and assembly factors. (**B**) Crown view of state A highlighting a protein-protein interaction hub with color-coded assembly factors.

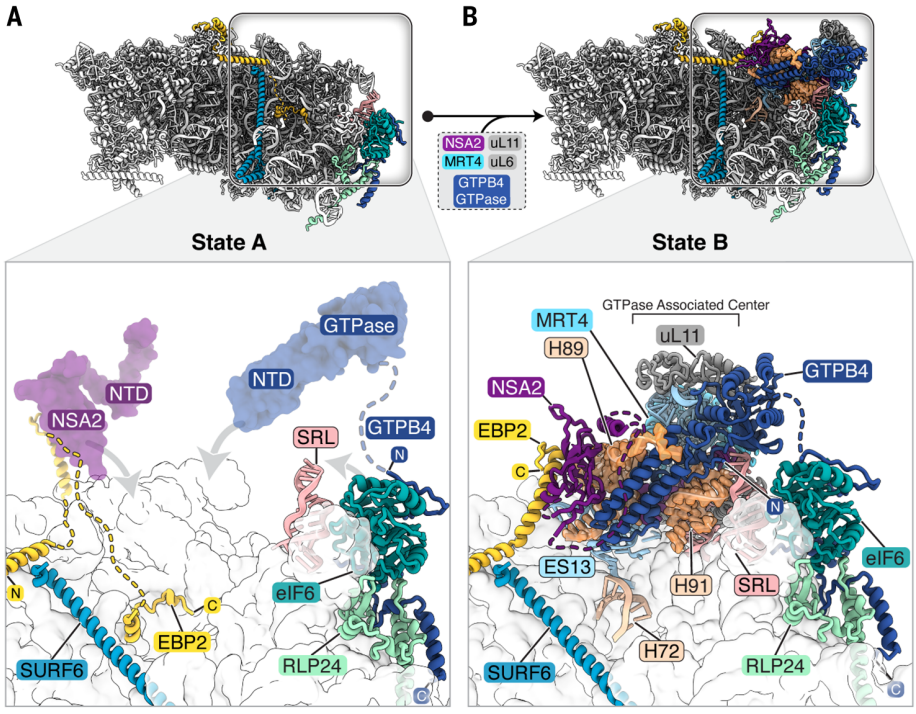



**Fig. 3. GTPB4 drives early PTC installation.** (**A** and **B**) Crown views of states A (A) and B (B) showing selected assembly factors. Bottom panels show a magnified view in which arrows show flexible elements of GTPB4, NSA2, and EBP2 in state A (A) before these are incorporated in state B (B). Assembly factors are color coded, and newly integrated PTC rRNA is represented as an orange surface.

remains GDP bound in all subsequent nucleolar and nuclear states that we observed (Figs. 1 and 3B and figs. S23 and S24). Because state B is the first state in which parts of the nascent PTC (helices 89 and 91) and the GAC (remodeled ES13 and helices 42 to 44) appear, these data indicate that the incorporation of these RNA elements is coupled to the GTPase activity of GTPB4. The incorporation of GTPB4 after GTP hydrolysis is therefore similar to other enzymes with ATPase or methyltransferase activity (*10*, *35*). Consistent with this model, conserved N-terminal residues of NSA2 that have been implicated in activating GTPase activity in yeast (*36*) are positioned near the GTPase active site (fig. S23B), and dominant-negative mutations near the GTPB4 active site result in nucleolar accumulation of GTPB4/Nog1 in mouse, consistent with trapped particles equivalent to state A in which only the C-terminal fragment of GTPB4 is bound (*32*).

State B represents a key transition state, because early assembly factors bridging domains I, II, and VI (SURF6-SSF1-RRP15) are still present, whereas EBP2 has been remodeled to accommodate helix 72 and stabilize NSA2 and the GTPase domain of GTPB4 in the posthydrolysis state (GDP bound) (Figs. 3B and fig. S23B). In state B, crucial rRNA elements around GTPB4, such as expansion segment 13 (ES13) and parts of helix 89 (H89), adopt immature conformations compared with the subsequent state C, in which the L1 stalk is incorporated (fig. S23, C and D).

Together, these data allow us to postulate a universal mechanism of early GTPase-mediated installation of the PTC that rationalizes existing data from bacterial and eukaryotic systems. In contrast to the prevalent model, in which GTP hydrolysis triggers dissociation of GTPB4 and its homologs in yeast and *E. coli* during late stages of assembly (*33*, *34*), our model postulates that GTP hydrolysis is used to dock GTPB4 and its homologs early as a quality control step to ensure correct PTC installation. However, in the absence of a structure of an early GTP-bound state, a late nucleotide exchange followed by GTP hydrolysis cannot be excluded.

### DDX54 couples irreversible rRNA remodeling to PTC assembly

After the GTPase-mediated partial installation of the PTC, the presence of GTPB4 in state B provides a unique binding site for the conserved DEAD-box RNA helicase DDX54 (Dbp10 in yeast) that has been implicated in PTC maturation downstream of GTPB4 (*37*–*39*). In addition to the two characteristic RecA lobes shared by all DEAD-box RNA helicases, DDX54 contains a structured C-terminal region (Elbow), N- and C-terminal extensions (NTEs and CTEs), as well as a C-terminal tail, providing substrate specificity (*9*). We found that DDX54 is incorporated into state C in a posthydrolysis state and remains ADP bound in states C to F (Figs. 1 and 4 and fig. S24).

Our structures illustrate that DDX54 has a pivotal role during nucleolar large ribosomal subunit assembly by performing three essential functions. First, DDX54 is responsible for the formation of functional centers such as the PTC and the L1 stalk. Second, DDX54 provides directionality to 60*S* maturation through remodeling of the pre-rRNA root helices such that early assembly factors can dissociate. Third, DDX54 acts as a recruitment platform for new assembly factors that catalyze downstream events (Fig. 4 and fig. S24).

In its first role during the formation of functional centers, the association of DDX54 with GTPB4 limits conformational freedom of the remaining elements of the PTC (H90, H92, and H93) and RNA elements that accompany the NOP2/NIP7 complex, including the L1 stalk and part of the E site. By acting as a clamp, DDX54 is thus able to facilitate PTC formation, remaining bound on top of H90 and H92 and fixing the L1 stalk in an immature conformation, consistent with prior biochemical data (*38*) (Fig. 4, A and B, and fig. S24E). The installation of elements of the PTC is therefore catalyzed through checkpoints that are controlled by energy-consuming enzymes in two steps, with GTPB4 installing H89 and H91 and DDX54 stabilizing H90, H92, and H93.

In its second role, DDX54 remodels the root helices, first of domain IV, helix 61 (H61), and subsequently domain III (H47), triggering a series of irreversible conformational changes that provide directionality to 60*S* assembly (Fig. 4, A and B, and fig. S24D). In state B, root helices of domains III and IV are trapped by the SURF6-RRP15-SSF1 complex (Fig. 4A). In state C, DDX54 has successfully remodeled H61, stabilizing an extended conformation of H61 that abolishes the binding site for the SURF6-RRP15-SSF1 complex (Fig. 4B and fig. S24D). The unidirectionality of the removal of these proteins is further guaranteed by a FTSJ3 C-terminal segment that binds the site previously occupied by SURF6, thereby preventing reassociation of the state A- or B-specific SURF6-RRP15-SSF1 complex while additionally blocking the nascent polypeptide exit tunnel together with RPF1 (Fig. 4, A and B, and fig. S24D). As a result of the release of the SURF6-RRP15-SSF1 complex, which previously maintained domain VI in an immature conformation, a compaction occurs between domains V and VI, creating a binding site for LLPH at the solvent side (Fig. 4B). The remodeling of H61 further involves alternate base pairing in the extension of the root helix of domain IV, thereby repositioning H64 and accommodating the newly integrated PTC elements H90 and H92. The specific alternate base pairing on top of H61 is further interrogated by DDX54 through a

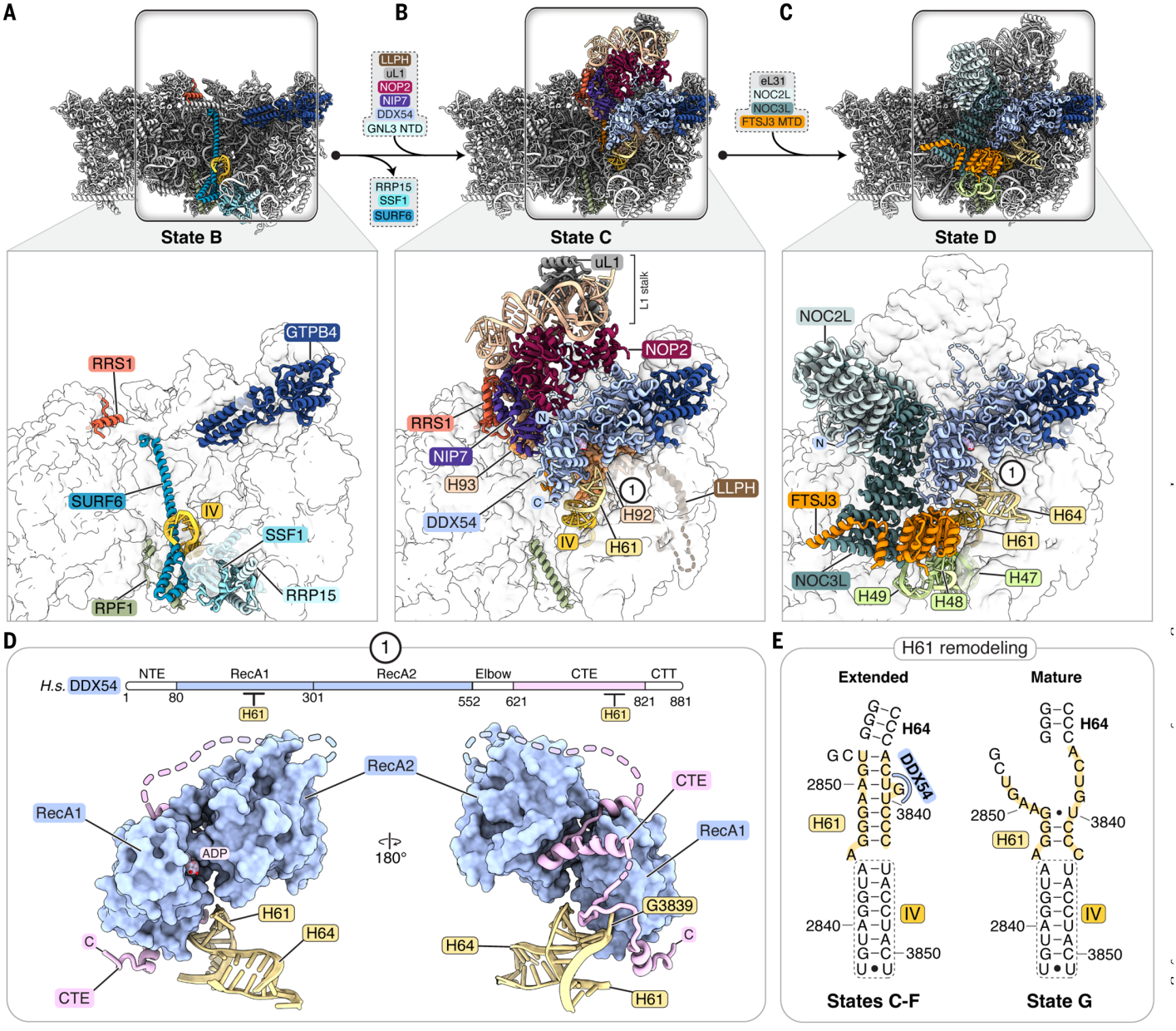



**Fig. 4. DDX54 integrates key transitions during early nucleolar assembly.** (**A** to **C**) Crown views of states B (A), C (B), and D (C) showing selected assembly factors as they enter the respective states. Magnified views below provide more detail of the environment present before (A), during (B), and after the installation of DDX54 (C). (**D**) Schematic indicating the domain organization of DDX54 and interactions with H61 of the 28*S* rRNA. The cartoon representation of DDX54 below shows two views of its interactions with a remodeled H61. A flipped-out base (G3839) is indicated. (**E**) Schematics of remodeled H61 as seen in states C to F (left) and the mature form of H61 as seen in state G and subsequent states (right).

flipped-out nucleotide (G3839) so that only the correct substrate can be accommodated near the PTC (Fig. 4, D and E, and fig. S24E).

In its third role as a recruitment platform, the C- and N-terminal extensions of DDX54 selectively recruit NOP2-NIP7 in state C and the FTSJ3-NOC2L-NOC3L assembly factor complex in state D, respectively (Fig. 4, B and C, and fig. S24, E and F). In state D, the NOC2L-NOC3L complex is positioned in a downward orientation, thereby preventing premature maturation of domain III (Figs. 1 and 4C).

The ability of DDX54 to interact with NOC2L-NOC3L and the methyltransferase domain of FTSJ3 through the DDX54 C-terminal tail allows DDX54 to organize elements of domain III that include the domain III root helix H47 as well as helices H48 and H49 (Fig. 4C and fig. S24, E and F). In this context, the methyltransferase domain of FTSJ3 performs a secondary structural role, whereas its substrate for 2′-O-methylation (nucleotide 4499 within the A-loop/H92) is already deeply buried underneath DDX54 (fig. S24E). The ability to tether FTSJ3 already in state A (Fig. 2B) is compatible with the modification of nucleotide 4499 before state G when FTSJ3 is removed.

The structural changes observed during the transitions from state B to D highlight how an ensemble of factors centered around the DEAD-box RNA helicase DDX54 remodels important segments of pre-rRNA. Beyond enzymes such as DDX54 and FTSJ3, which act directly upon a substrate, these also include proteins that fulfill either architectural roles such as NOC2L-NOC3L or stabilizing roles such as the N terminus of

GNL3 that stabilizes ES13 in state C specifically (Fig. 4 and fig. S23C). These findings support a model for the cooperative actions of these factors that is required to advance 60*S* assembly, further rationalizing prior genetic interactions observed in yeast among Dbp10, Noc3, and Nug1 (*39*) (Fig. 4).

### Transitions toward nuclear maturation

After a rotation of the NOC2L-NOC3L complex from a downward (state D) into an upward position that is accompanied by the removal of MAK16-assembly factors at the solvent-exposed side (state E; Fig. 1 and fig. S25), we observed that a trio of WD40 repeat proteins, BOP1, WDR12, and WDR55 (yeast Erb1, Ytm1, and Jip5) chaperones the formation of domain III in state F (figs. S25 and S26). Subsequently, the transition from state F to G is marked by the MDN1-mediated removal of a nucleolar protein interaction network. In state G, DDX54 is replaced by a peptide of the GTPase NOG2 that binds to GTPB4 (Fig. 1 and fig. S27). The initial tethering of NOG2 is followed by the installation of the 5*S* RNP in a nonrotated state in state H, for which we have obtained a partial reconstruction (Fig. 1 and fig. S27). For a more detailed description of these transitions, see the supplementary text.

### The rixosome and NOP53 couple 5*S* RNP rotation to ITS2 processing

After nucleolar assembly of the human pre-60*S*, a critical event during the subsequent nuclear maturation of pre-60*S* precursors is the MDN1-catalyzed rotation of the 5*S* RNP from the unrotated to the rotated state, which results in the installation of the CP (*7*, *15*, *16*, *18*, *40*–*43*).

The structure of state I shows that the formation of this intermediate, in which the 5*S* RNP is rotated, is the product of several hierarchical and coupled steps (Fig. 5). Only upon correct rotation of the 5*S* RNP and the departure of early assembly factors (RPF2 and RRS1) can the assembly factor SDA1 and the rixosome stably associate, thereby stabilizing the L1 stalk in a bent conformation (Fig. 5). Association of the rixosome is required for ITS2 processing on two levels: The rixosome-associated endonuclease LAS1L (*16*, *22*) initiates cleavage of ITS2, whereas the bent conformation of the L1 stalk, induced by rixosome stabilization, serves as a binding site for a peptide of NOP53 that brings the RNA exosome into close proximity to its ITS2 substrate (Fig. 5 and fig. S28). Although the RNA exosome and some components of the rixosome (MDN1, SENP3, and NOL9) could not be visualized in our reconstructions, most likely because of flexibility, they are present in our mass spectrometry data (data S1).

Although 5*S* RNP rotation and bending of the L1 stalk have been observed in the yeast system, limiting resolution near ITS2 has so far precluded a mechanistic understanding of how these large-scale conformational changes are coupled to processing of ITS2 by the RNA exosome (*7*, *17*, *40*). Our high-resolution structures of state I explain how the bent conformation of the L1 stalk and the presence of ITS2 are specifically interrogated by 200 amino acids of a universally conserved N-terminal segment of NOP53 (sensor and anchor domains), which had not previously been visualized (Fig. 5B and fig. S28). The involvement of the NOP53 sensor domain is supported by biochemical data showing that N-terminal truncations of this domain result in ITS2-processing defects (*44*).

The NOP53 sensor domain is only able to bind the L1 stalk in the bent conformation, binding in such a way that in the presence of ITS2 and associated assembly factors, the NOP53 anchor domain can additionally wrap around ITS2. The joint binding of the sensor and anchor domains positions a recruitment peptide of the RNA exosome, the NOP53 AIM motif (*45*), in very close proximity to ITS2. Because the AIM motif has been shown to interact with the RNA exosome–associated RNA helicase MTR4 (*46*, *47*), our structural data now provide key insights into the coupling of 5*S* RNP rotation, L1 stalk conformation, and ITS2 processing.

The structure of state I now rationalizes how defects or depletions of assembly factors involved in 5*S* RNP rotation and the installation of the rixosome result in the accumulation of unprocessed pre-rRNA species (*18*–*22*). In the absence of a stabilized bent L1 stalk, either no cleavage of ITS2 occurs or in cases where ITS2 cleavage does occur, the ITS2 precursor cannot be efficiently processed by the RNA exosome because it is not positioned close to its substrate by NOP53 (*28*). This model also explains how the NOP53 anchor domain can remain associated with the ITS2 region during exosome-mediated ITS2 processing (*17*).

### Functional integration of the human rixosome

State I further illustrates how the human rixosome and NOG2 function through an evolutionarily expanded protein-protein interaction network that is formed through CTEs of the assembly factors NOG2, TEX10, and LAS1L, which have no counterparts in yeast (Fig. 5A and fig. S29). These CTEs facilitate three integrated functions during nuclear pre-60*S* maturation.

First, extensive interactions of CTEs with NOG2 allow the human rixosome to already associate with late nucleolar assembly through a peptide of NOG2, as revealed in states G and H (Fig. 1 and fig. S27) (*10*). Second, the CTE of NOG2 is also ensuring directionality during the

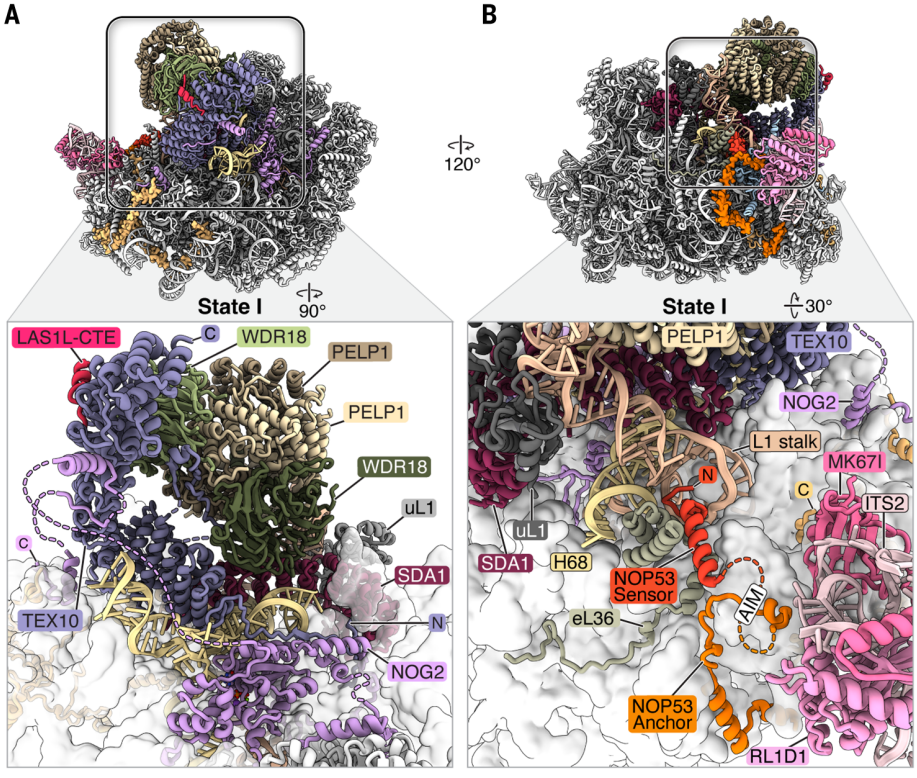



**Fig. 5. State-specific processing of ITS2.** (**A** and **B**) Two views of state I showing selected assembly factors involved in ITS2 processing. Bottom panels show a magnified view of the rixosome and its interactions with the pre-60*S* particle (A) and the NOP53 N terminus near the L1 stalk (B). The NOP53 sensor and anchor domains (red and orange, respectively) are shown in proximity to the AIM motif.

transition from nucleolar to nuclear assembly by probing the environment around eL34, where it has a mutually exclusive binding site with the nucleolar assembly factor FTSJ3 (yeast Spb1) (Fig. 5A and figs. S22B and S29D). Third, the large C-terminal extension of TEX10 provides binding sites for NOG2, as well as several rixosome components known as the RIX1 subcomplex (comprising two copies each of PELP1 and WDR18) (*48*) and the C-terminal extension of the endonuclease LAS1L, thereby directly recruiting the ITS2 endonuclease to state I (Fig. 5A and fig. S29D). In the context of the human rixosome, CTEs are used to flexibly tether different enzymatic activities, with LAS1L, NOL9, and SENP3 acting as functional modules with endonuclease, 5′ polynucleotide kinase, and SUMO-specific protease activity, respectively (fig. S29, C and D) (*27*, *48*, *49*).

Together, our structural data on state I highlight specific aspects of human ribosome assembly in which the rixosome has evolved into a higher-order protein interaction hub that directly couples 5*S* RNP rotation and L1 stalk bending to the recruitment of the endonuclease LASL1 to initiate cleavage of ITS2 for subsequent exosome-mediated RNA processing (Fig. 5 and figs. S28 and S29).

### Processing of recombinant pre-rRNA in human cells

Human ITS2 is five-fold larger than that of yeast. Although available biochemical data suggest that there are more pre-rRNA–processing sites in human cells (*50*), it remains unclear how ITS2 processing is achieved and which of its RNA elements are essential for cleavage and ribosome assembly. Similarly, it is unclear whether large RNA expansion segments of the human 28*S* rRNA are also required for ribosome assembly. By combining our structural data with a human large ribosomal subunit assembly assay, we have determined the order of ITS2 removal and identified key regions of pre-rRNA that are required for LSU maturation (Fig. 6).

Similar to what has been observed for the 5′ external transcribed spacer in the context of the human SSU processome (*51*), only a very small portion of ITS2 can be visualized in all pre-60*S* assembly intermediates (fig. S30). In state I1, only helices 1 and 2 of ITS2 are visible and are positioned close to the ITS2-associated assembly factors MK67I, RLP7, PES1, and RL1D1 (in yeast Nop15, Rlp7, Nop7, and Cic1, respectively) (Fig. 6A and fig. S30). States I1 to I3 represent different stages of ITS2 processing and disassembly of the associated protein factors, with elements of ITS2, RL1D1, and the NOP53 anchor domain disappearing in state I2 (Fig. 6B). The transition from state I2 to I3 is marked by the disordering of the remaining N-terminal segments of NOP53 (sensor and anchor domains) and of the structured domain of RLP7 that is associated with MK67I (Fig. 6, B and C).

The minimal visualized region of ITS2 within the pre-60*S* particles prompted us to investigate whether disordered elements of ITS2 and expansion segments within human 28*S* rRNA are required for ribosome assembly and pre-rRNA processing by designing truncations in these regions and testing for successful ribosome assembly (Fig. 6, D to F, and fig. S30, B and C). Whereas truncations of ITS2 were based on its predicted secondary structure (fig. S30B), 28*S* rRNA truncations were focused on large expansion segments that were not visualized in any of our human pre-60*S* assembly intermediates (Fig. 6D). All truncations were inserted into a plasmid containing an engineered human rDNA that, in addition to the indicated truncations, includes unique 28*S* and ITS2 probes for Northern blotting, specifically detecting the formation of mature recombinant large ribosomal subunits, as well as its ITS2-containing precursors (Fig. 6D and fig. S30, B and C).

Within ITS2, long helices were selected for truncations that either lacked known cleavage sites (Δ1 and Δ2) or contained the predicted LAS1L cleavage site responsible for generating a 12*S* species (*52*) (Δ3), as well as combinations of these (Δ1,2 and Δ1,2,3). We further designed ITS2 variants that lacked the LAS1L cleavage site and contained RNA segments just exceeding what is visible in our structures (mini and micro; Fig. 6D and fig. S30, B and C). Within the 28*S*, rRNA truncations were designed to eliminate disordered expansion segments ES7ab and ES15 (Δa), ES27 (Δb), or all three of these (Δa,b). The resulting truncations reduced ITS2 from five times the length of yeast ITS2 to ~60% of the length. Similarly, the combined truncations of expansion segments of the 28*S* rRNA resulted in a 28*S* rRNA that is only 20% larger than its yeast counterpart (Fig. 6E).

To investigate the effects of these truncations, the designed variants were transfected into human embryonic kidney (HEK)-293F cells from which cytoplasmic and nuclear extracts were obtained to identify engineered mature ribosomes that were exported into the cytoplasm, as well as their corresponding nucleolar and nuclear precursors (Fig. 6F and fig. S31). We found that ITS2 truncations that removed long helices without the predicted LAS1L cleavage site (Δ1, Δ2, and Δ1,2) were processed similarly to the wild type without accumulation of a nucleolar precursor species (32*S*). By contrast, the elimination of the predicted LAS1L cleavage site (Δ3 or Δ1,2,3) resulted in the accumulation of a 32*S* species in the nucleus. Despite this, mature ribosomes containing either 28*S* or 32*S* pre-rRNA were detected in the cytoplasm for Δ3 or Δ1,2,3 truncations. These data suggest that either redundant enzymes or cleavage sites exist to trigger ITS2 removal, and that despite the absence of a regular LAS1L cleavage site, some precursors are exported. These observations are consistent with what has been observed in yeast in the absence of Las1, where ribosomes containing ITS2 are exported into the cytoplasm and turned over (*53*–*55*) (Fig. 6F). The strongest accumulation of precursors was observed with minimal variants of ITS2 (mini and micro truncations), for which no processing of ITS2 was observed and the resulting cytoplasmic ribosomes contained ITS2 (Fig. 6F). For truncations of expansion segments within the 28*S* pre-rRNA, no processing defects were observed compared with the wild type (Fig. 6F). Therefore, large parts of ITS2 and the 28*S* expansion segments ES7ab, ES15, and ES27 are dispensable for ribosome assembly in *cis* in human cells, with the exception of the regions of ITS2 including the pre-60*S* associated helices and the LAS1L cleavage site, the removal of which results in partial processing defects.

### Transitions preceding nuclear export

After the departure of the human rixosome, large-scale RNA conformational changes, together with eukaryote-specific rRNA and protein elements, drive the maturation of the central protuberance and the formation of the human E site (figs. S32 and S33). Subsequently, we observed states in which the late nuclear assembly factor TMA16, together with the newly identified assembly factor L10K, triggered the removal of PTC-proximal assembly factors such as GTPB4, NOG2, and NSA2 (fig. S34). For a more detailed description of these states, see the supplementary text.

### Chemical modifications of pre-RNA during human pre-60*S* assembly

Beyond our insights into the role of numerous human nuclear ribosome assembly factors, the high resolution of our cryo-EM reconstructions of nuclear pre-60*S* assembly intermediates (states I to L) also allowed us to introduce and confirm the positions of many previously mapped covalent modifications of the human 28*S* rRNA (*56*, *57*). These include modifications introduced through proteins that are guided by small nucleolar RNAs (snoRNAs) base pairing with RNA regions surrounding their substrates (2′-O-methylations and pseudouridylations), as well as modifications introduced through protein factors directly ($m^1A$, $m^3U$, $m^5C$, and $m^6A$ nucleobase modifications and a 2′-*O*-methylation) (*58*–*63*) (fig. S35). During the assembly of the human pre-60*S*, the introduction of every chemical modification will require local chaperoning of substrate RNA, so any introduced modification is a direct representation of a prior RNA chaperoning event. By mapping the observed chemical modifications onto key nucleolar (states A to H) and nuclear assembly intermediates (states I to L) of our human pre-60*S* structures, we could now observe the order in which RNA elements appear and place chemical modifications

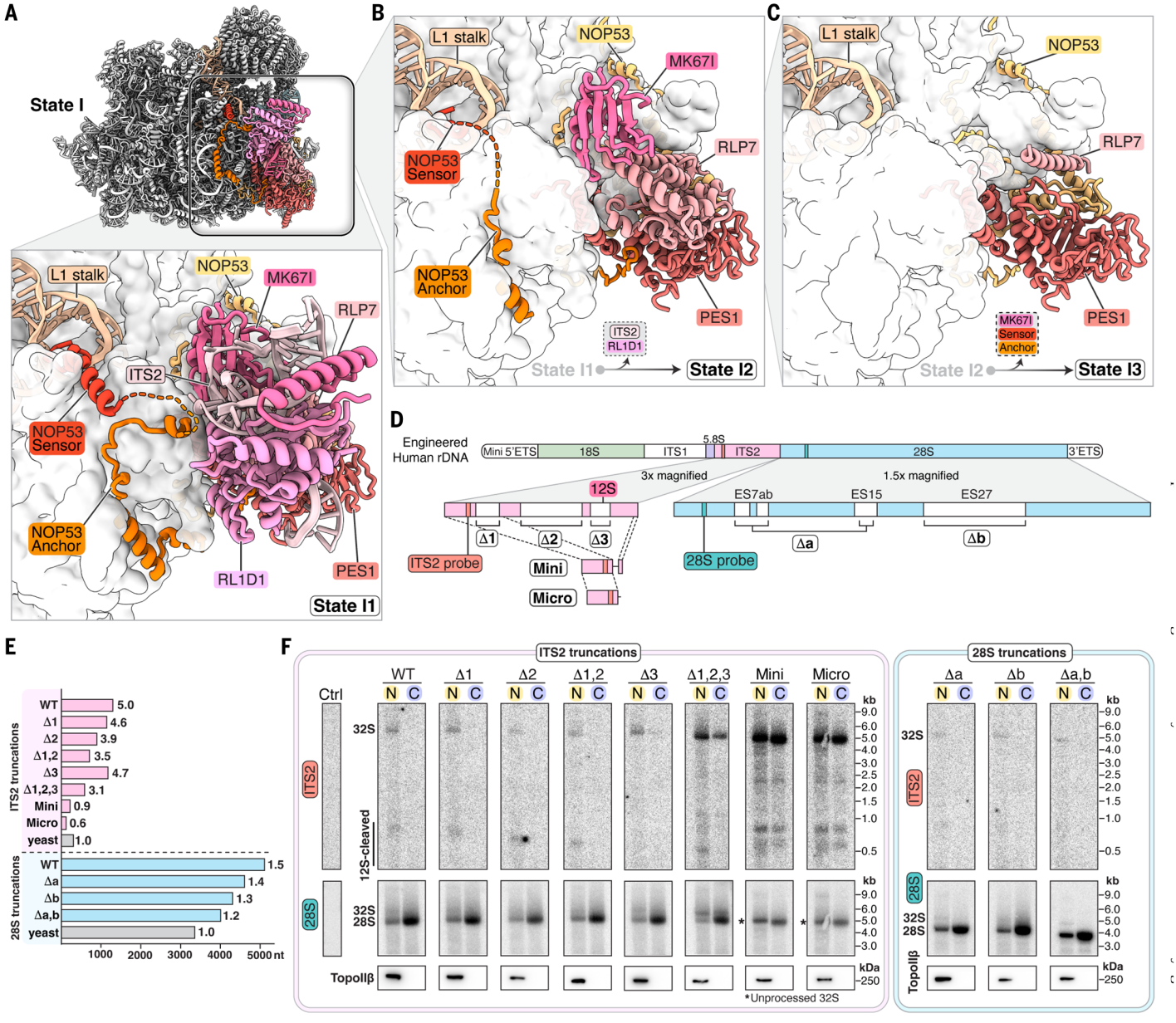



**Fig. 6. Processing of synthetic human 28*S* in vivo.** (**A**) View of state I with color-coded assembly factors near ITS2. The magnified view below shows domains of NOP53 around ITS2-associated assembly factors and ITS2 (light pink). (**B** and **C**) Comparative magnified views of the ITS2 region in substates I2 and I3. (**D**) Schematic representation of designed ITS2 variants in relation to an engineered human rDNA locus. Magnified sections indicate the position of ITS2 variants (bottom left) and designed truncations of expansion segments ES7ab and ES15 (Δa) or ES27 (Δb) (bottom right). The locations of specific probes for ITS2 and 28*S* are highlighted in light red and cyan, respectively. (**E**) Bar graph indicating the size of human ITS2 (pink, top) and 28*S* (light blue) variants relative to their yeast counterparts (gray). (**F**) Northern blots of indicated synthetic human ribosomes containing truncations in ITS2 (left) and 28*S* (right) using probes for ITS2 (red, top) and 28*S* (cyan, bottom) for both nuclear (N, yellow) and cytoplasmic (C, blue) fractions. Western blots for TopoIIβ highlight the separation of nuclear and cytoplasmic extracts (bottom).

in the context of ribosome assembly (Fig. 7). This analysis highlights four central concepts for human pre-*60S* assembly, as described below.

First, during nucleolar stages, domains I to VI assemble in an order similar to what has been previously observed in yeast (*11*–*13*), with domains I, II, and VI folding first before domains V, III, and IV are integrated (Fig. 7A, fig. S36, and data S2). Second, already in state A, all root helices, which are the origin of the six 28*S* rRNA subdomains, are visible. (fig. S20). A high density of chemical modifications is observed near or within these root helices, suggesting important prior RNA chaperoning and thus highly controlled RNA folding before the formation of state A (Fig. 7, A and B). Third, consistent with the notion that the presence of a chemical modification requires prior RNA chaperoning, we found that regions that are heavily modified form late during assembly, such as elements of domains V and IV. These regions include the PTC that is installed in two steps through the GTPase GTPB4 in state B (PTC-I; Fig. 7B) and the DEAD-box RNA helicase DDX54 in state C (PTC-II; Fig. 7B). In the context of ribosome assembly, the presence of a modified nucleotide can therefore be seen as a means of either delaying folding in a particular region or preventing misfolding. This logic further implies that

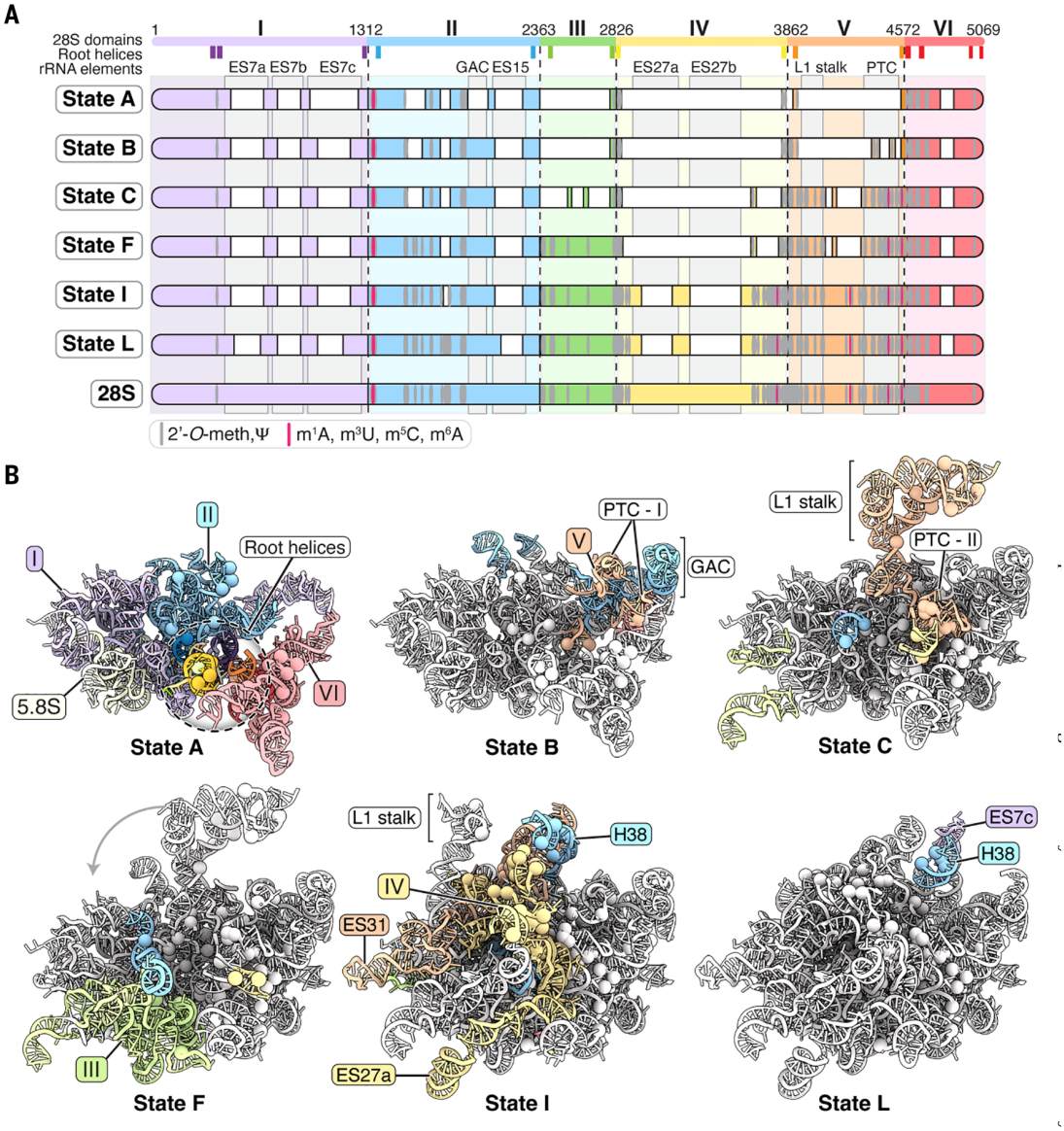



**Fig. 7. Chemical modifications of human 28*S* rRNA during assembly.** (**A**) Schematic domain organization of the six color-coded domains (I to VI) of the human 28*S* rRNA. Root helices are indicated as darker blocks. Below the schematic, visible RNA elements are indicated for different states of the human pre-60*S* assembly pathway in the nucleolus (states A, B, C, and F) and nucleus (states I and L). Chemical modifications are shown in gray or red as indicated. The positions of selected functional centers (PTC, GAC, and L1 stalk) and RNA expansion segments (ES7, ES15, and ES27) are shown. (**B**) RNA structures of assembly intermediates shown in (A) with newly folded RNA elements color coded. Root helices and functional centers (GAC, PTC, and L1 stalk) are indicated. Chemical modifications are represented as spheres.

regions that need to fold quickly at the beginning of assembly, such as domain I, should have a low density of chemical modifications, which is indeed what we observed. Fourth, regions that are not conserved, disordered in our reconstructions, or not subject to extensive RNA remodeling should lack chemical modifications. This is precisely what we observed for expansion segments such as ES7, ES15, and ES27, which our functional studies show to be dispensable for ribosome assembly (Figs. 6 and 7A).

The mapping of human chemical RNA modifications as a function of ribosome assembly in the nucleolus and nucleus provides a foundation with which to study both universally conserved chemical modifications and species-specific adaptations.

## Discussion

During the nucleolar and nuclear stages of pre-60*S* assembly, functional centers need to be first assembled and subsequently matured. On the basis of our structural data, we now propose a simplified model for human nucleolar and nuclear pre-60*S* biogenesis (Fig. 8) that complements all prior work in *S. cerevisiae* (fig. S37). We highlight that emerging key nucleolar themes include high-level organization and hierarchy at the pre-rRNA and protein level (Figs. 2 and 8), the use of energy-consuming enzymes such as GTPB4 and DDX54 to drive transitions, and the installation of functional centers in a unidirectional manner (Figs. 3, 4, and 8). Critical nuclear themes include recognition of pre-rRNA for LAS1L-mediated cleavage, followed by RNA exosome–mediated processing of ITS2 and how the human rixosome can integrate these processing steps through C-terminal protein extensions (Figs. 5 and 8).

Beyond representing a near-complete structural inventory of nucleolar and nuclear human pre-60*S* assembly, our data now start to interweave human ribosome assembly with research fields that have so far appeared disconnected at a mechanistic level. The observation of posttranslational modifications of ribosome assembly factors as observed for BOP1 (Fig. 2 and fig. S22B) suggest a direct mechanistic role in the molecular control of human ribosome assembly and indicates how signaling pathways may integrate human ribosome assembly with cellular metabolism. In addition, our work lays the foundation for a mechanistic understanding of human diseases so that the molecular effects of point mutations in ribosomal proteins can be put into a temporal context of large ribosomal subunit assembly in patients suffering from a range of ribosomopathies (*64*–*66*).

At the nuclear level, the structural basis for the coupling between 5*S* RNP rotation and ITS2 processing, as shown in state I (Fig. 5), highlights the importance of structural studies

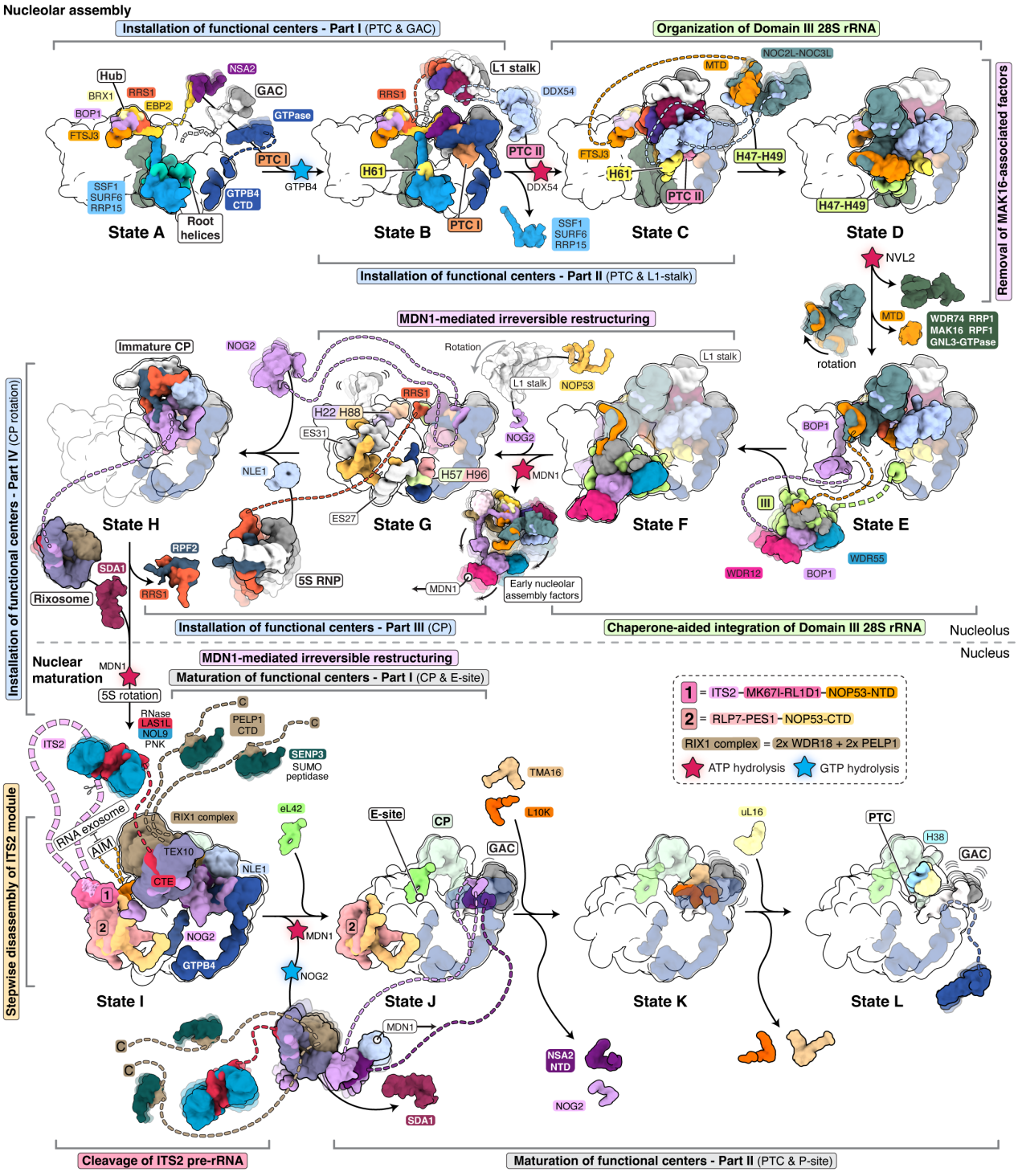



**Fig. 8. Model for human pre-60*S* biogenesis in the nucleolus and nucleus.** Simplified schematic model of human large ribosomal subunit biogenesis in the nucleolus (top) and nucleus (bottom). Key nucleolar transitions such as the installation of functional centers (PTC, GAC, and L1 stalk, and CP; blue highlights), the domain III integration (green highlights), and the irreversible steps involving protein removal (pink highlights) are indicated. In the nucleus, critical maturation steps such as ITS2 cleavage by LAS1L (red highlight), subsequent stepwise disassembly of ITS2 module by the RNA exosome (orange highlight), and maturation of functional centers (CP, E site, PTC, and P site; gray highlight) are indicated. Irreversible GTP and ATP hydrolysis events are indicated with blue and red stars, respectively. Color-coded tethered protein and RNA elements are indicated with linkers and motion blur.

with endogenous complexes. Further, it underlines the complexity of eukaryotic ribosome assembly in which MDN1-induced *5S* RNP rotation precedes ITS2 processing with Las1 cleavage and Nop53-mediated recruitment of the RNA exosome, events that cannot be reconstituted by using only enzymes involved in ITS2 processing (*17*, *67*).

Across nucleolar and nuclear assembly, our analysis of chemical modifications of RNA indicates that chemical modifications can be used as a readout of prior RNA chaperoning

during the assembly of human pre-60*S* subunits (Fig. 7). This notion is in agreement with additional functions that have been described for chemical modifications. First, chemical modifications in yeast and human cells may also serve as checkpoints during assembly, where they may be interrogated by assembly factors (fig. S29B) (*10*). Second, although chemical modifications can result in additional stabilization of the ribosome, complete ribosomes can be formed even in the presence of catalytically dead enzymes involved in pseudouridylation while the conformational freedom of subunits including intersubunit rotation is partially affected (*68*). Third, the heterogeneity of chemical modifications as observed in the human ribosome (*69*) can be explained as the result of several redundant sites of chemical modifications that together guarantee efficient ribosome assembly without the need for selective and quantitative modification. These data suggest that in archaeal and eukaryotic systems, in parallel to proteinaceous assembly factors, snoRNA-mediated modifications have evolved as a second layer of control to ensure the productive biogenesis of ribosomes.

## Methods summary

Full details of the material and methods can be found in the supplementary materials. In brief, CRISPR/Cas9-based genome editing was used to generate a biallelically tagged MK67I-GFP cell line from which human pre-60*S* assembly intermediates were affinity purified and structurally characterized using cryo-EM. In addition, an in vivo recombinant ribosome assembly assay using an engineered human rDNA locus was used to investigate whether rRNA elements within ITS2 and the 28S rRNA are required for the large ribosomal subunit biogenesis.

### *MK67I-GFP$^{++}$ cell line generation*

Two DNA repair templates were designed, each encoding a GFP tag followed by a surface display sequence. These templates were cloned into a vector along with a plasmid expressing the single-guide RNA and the Cas9 enzyme. The DNA mixture was transfected into HEK-293F cells, which were then allowed to recover and expand. Cells expressing GFP were selected using fluorescence-assisted cell sorting, and then cells positive for both surface display epitopes were further selected. Single-cell clones were expanded and genotyped to confirm biallelic editing.

### *Purification of human pre-60S assembly intermediates*

For the purification of human pre-60*S* assembly intermediates, the MK67I-GFP cell line was grown in suspension cultures. Cells were harvested and nuclei were extracted from the cells. Intact nuclei were resuspended in solubilization buffer, and the lysate was spun down to remove the insoluble fraction. The supernatant was incubated with anti-GFP nanobody beads. The complexes were eluted using 3C-protease cleavage. Purified pre-60*S* particles were used for further analysis.

### *Recombinant human rDNA engineering*

A full copy of the human rDNA sequence containing different truncations in ITS2 and 28S was cloned into a plasmid, and unique probe sequences were introduced to monitor the production of recombinant ribosomes. The engineered rDNA plasmids were transfected into human HEK-293F cells, which were then used for isolation of recombinant ribosomes from different cellular compartments and further analysis by Northern blot.

### *Cryo-EM data collection and processing*

Cryo-EM grids were prepared using the purified pre-60*S* particles. Cryo-EM images were collected on a Titan Krios electron microscope. Cryo-EM data processing and model building were performed using standard software and procedures.

## REFERENCES AND NOTES


1. K. Scherrer, J. E. Darnell, Sedimentation characteristics of rapidly labelled RNA from HeLa cells. *Biochem. Biophys. Res. Commun.* **7**, 486–490 (1962). doi: 10.1016/0006-291X(62)90341-8; pmid: 14498283
2. L. Tafforeau *et al*., The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol. Cell* **51**, 539–551 (2013). doi: 10.1016/j.molcel.2013.08.011; pmid: 23973377
3. K. E. Bohnsack, M. T. Bohnsack, Uncovering the assembly pathway of human ribosomes and its emerging links to disease. *EMBO J.* **38**, e100278 (2019). doi: 10.15252/embj.2018100278; pmid: 31268599
4. S. Klinge, J. L. Woolford Jr., Ribosome assembly coming into focus. *Nat. Rev. Mol. Cell Biol.* **20**, 116–131 (2019). doi: 10.1038/s41580-018-0078-y; pmid: 30467428
5. V. Kargas *et al*., Mechanism of completion of peptidyltransferase centre assembly in eukaryotes. *eLife* **8**, e44904 (2019). doi: 10.7554/eLife.44904; pmid: 31115337
6. Y. Zhou, S. Musalgaonkar, A. W. Johnson, D. W. Taylor, Tightly-orchestrated rearrangements govern catalytic center assembly of the ribosome. *Nat. Commun.* **10**, 958 (2019). doi: 10.1038/s41467-019-08880-0; pmid: 30814529
7. L. Kater *et al*., Construction of the central protuberance and L1 stalk during 60S subunit biogenesis. *Mol. Cell* **79**, 615–628.e5 (2020). doi: 10.1016/j.molcel.2020.06.032; pmid: 32668200
8. M. Prattes *et al*., Visualizing maturation factor extraction from the nascent ribosome by the AAA-ATPase Drg1. *Nat. Struct. Mol. Biol.* **29**, 942–953 (2022). doi: 10.1038/s41594-022-00832-5; pmid: 36097293
9. V. E. Cruz *et al*., Sequence-specific remodeling of a topologically complex RNP substrate by Spb4. *Nat. Struct. Mol. Biol.* **29**, 1228–1238 (2022). doi: 10.1038/s41594-022-00874-9; pmid: 36482249
10. J. N. Yelland, J. P. K. Bravo, J. J. Black, D. W. Taylor, A. W. Johnson, A single 2′-O-methylation of ribosomal RNA gates assembly of a functional ribosome. *Nat. Struct. Mol. Biol.* **30**, 91–98 (2023). doi: 10.1038/s41594-022-00891-8; pmid: 36536102
11. L. Kater *et al*., Visualizing the assembly pathway of nucleolar pre-60S ribosomes. *Cell* **171**, 1599–1610.e14 (2017). doi: 10.1016/j.cell.2017.11.039; pmid: 29245012
12. Z. A. Sanghai *et al*., Modular assembly of the nucleolar pre-60S ribosomal subunit. *Nature* **556**, 126–129 (2018). doi: 10.1038/nature26156; pmid: 29512650
13. D. Zhou *et al*., Cryo-EM structure of an early precursor of large ribosomal subunit reveals a half-assembled intermediate. *Protein Cell* **10**, 120–130 (2018). doi: 10.1007/s13238-018-0526-7; pmid: 29557065
14. R. Martin, A. U. Straub, C. Doebele, M. T. Bohnsack, DExD/H-box RNA helicases in ribosome biogenesis. *RNA Biol.* **10**, 4–18 (2013). doi: 10.4161/rna.21879; pmid: 22922795
15. M. Prattes, Y.-H. Lo, H. Bergler, R. E. Stanley, Shaping the nascent ribosome: AAA-ATPases in eukaryotic ribosome biogenesis. *Biomolecules* **9**, 715 (2019). doi: 10.3390/biom9110715; pmid: 31703473
16. S. Wu *et al*., Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature* **534**, 133–137 (2016). doi: 10.1038/nature17942; pmid: 27251291
17. J. M. Schuller, S. Falk, L. Fromm, E. Hurt, E. Conti, Structure of the nuclear exosome captured on a maturing preribosome. *Science* **360**, 219–222 (2018). doi: 10.1126/science.aar5428; pmid: 29519915
18. J. Micic *et al*., Coupling of 5S RNP rotation with maturation of functional centers during large ribosomal subunit assembly. *Nat. Commun.* **11**, 3751 (2020). doi: 10.1038/s41467-020-17534-5; pmid: 32719344
19. T. I. Moy, D. Boettner, J. C. Rhodes, P. A. Silver, D. S. Askew, Identification of a role for Saccharomyces cerevisiae Cgr1p in pre-rRNA processing and 60S ribosome subunit synthesis. *Microbiology (Reading)* **148**, 1081–1090 (2002). doi: 10.1099/00221287-148-4-1081; pmid: 11932453
20. L. P. P. Cepeda *et al*., The ribosome assembly factor Nop53 controls association of the RNA exosome with pre-60S particles in yeast. *J. Biol. Chem.* **294**, 19365–19380 (2019). doi: 10.1074/jbc.RA119.010193; pmid: 31662437
21. K. Galani, T. A. Nissan, E. Petfalski, D. Tollervey, E. Hurt, Rea1, a dynein-related nuclear AAA-ATPase, is involved in late rRNA processing and nuclear export of 60 S subunits. *J. Biol. Chem.* **279**, 55411–55418 (2004). doi: 10.1074/jbc.M406876200; pmid: 15528184
22. C. Saveanu *et al*., Nog2p, a putative GTPase associated with pre-60S subunits and required for late 60S maturation steps. *EMBO J.* **20**, 6475–6484 (2001). doi: 10.1093/emboj/20.22.6475; pmid: 11707418
23. S. Schillewaert, L. Wacheul, F. Lhomme, D. L. J. Lafontaine, The evolutionarily conserved protein Las1 is required for pre-rRNA processing at both ends of ITS2. *Mol. Cell. Biol.* **32**, 430–444 (2012). doi: 10.1128/MCB.06019-11; pmid: 22083961
24. P. Mitchell, E. Petfalski, A. Shevchenko, M. Mann, D. Tollervey, The exosome: A conserved eukaryotic RNA processing complex containing multiple 3′—>5′ exoribonucleases. *Cell* **91**, 457–466 (1997). doi: 10.1016/S0092-8674(00)80432-8; pmid: 9390555
25. X. Liang *et al*., Structural snapshots of human pre-60S ribosomal particles before and after nuclear export. *Nat. Commun.* **11**, 3542 (2020). doi: 10.1038/s41467-020-17237-x; pmid: 32669547
26. H. Zhou *et al*., Rixosomal RNA degradation contributes to silencing of Polycomb target genes. *Nature* **604**, 167–174 (2022). doi: 10.1038/s41586-022-04598-0; pmid: 35355014
27. C. D. Castle, E. K. Cassimere, C. Denicourt, LAS1L interacts with the mammalian Rix1 complex to regulate ribosome biogenesis. *Mol. Biol. Cell* **23**, 716–728 (2012). doi: 10.1091/mbc.e11-06-0530; pmid: 22190735
28. J. Talkish, J. Zhang, J. Jakovljevic, E. W. Horsey, J. L. Woolford Jr., Hierarchical recruitment into nascent ribosomes of assembly factors required for 27SB pre-rRNA processing in Saccharomyces cerevisiae. *Nucleic Acids Res.* **40**, 8646–8661 (2012). doi: 10.1093/nar/gks609; pmid: 22735702
29. I. C. Kos-Braun, I. Jung, M. Koš, Tor1 and CK2 kinases control a switch between alternative ribosome biogenesis pathways in a growth-dependent manner. *PLOS Biol.* **15**, e2000245 (2017). doi: 10.1371/journal.pbio.2000245; pmid: 28282370
30. P. V. Hornbeck *et al*., PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43** (D1), D512–D520 (2015). doi: 10.1093/nar/gku1267; pmid: 25514926
31. A. Vanden Broeck, S. Klinge, An emerging mechanism for the maturation of the small subunit processome. *Curr. Opin. Struct. Biol.* **73**, 102331 (2022). doi: 10.1016/j.sbi.2022.102331; pmid: 35176592
32. Y. R. Lapik, J. M. Misra, L. F. Lau, D. G. Pestov, Restricting conformational flexibility of the switch II region creates a dominant-inhibitory phenotype in Obg GTPase Nog1. *Mol. Cell. Biol.* **27**, 7735–7744 (2007). doi: 10.1128/MCB.01161-07; pmid: 17785438
33. P. Klingauf-Nerurkar *et al*., The GTPase Nog1 co-ordinates the assembly, maturation and quality control of distant ribosomal functional centers. *eLife* **9**, e52474 (2020). doi: 10.7554/eLife.52474; pmid: 31909713
34. R. Nikolay *et al*., Snapshots of native pre-50S ribosomes reveal a biogenesis factor network and evolutionary specialization.

*Mol. Cell* **81**, 1200–1215.e9 (2021). doi: 10.1016/j.molcel.2021.02.006; pmid: 33639093
35. V. E. Cruz *et al*., Sequence-directed RNA remodeling within a topologically complex RNP substrate. bioRxiv (2022), doi: 10.1101/2022.02.24.481786.
36. H. Paternoga *et al*., Mutational Analysis of the Nsa2 N-Terminus Reveals Its Essential Role in Ribosomal 60S Subunit Assembly. *Int. J. Mol. Sci.* **21**, 9108 (2020). doi: 10.3390/ijms21239108; pmid: 33266193
37. F. Burger, M. C. Daugeron, P. Linder, Dbp10p, a putative RNA helicase from Saccharomyces cerevisiae, is required for ribosome biogenesis. *Nucleic Acids Res.* **28**, 2315–2323 (2000). doi: 10.1093/nar/28.12.2315; pmid: 10871363
38. R.-G. Manikas, E. Thomson, M. Thoms, E. Hurt, The $K^+$-dependent GTPase Nug1 is implicated in the association of the helicase Dbp10 to the immature peptidyl transferase centre during ribosome maturation. *Nucleic Acids Res.* **44**, 1800–1812 (2016). doi: 10.1093/nar/gkw045; pmid: 26823502
39. J. Bassler, M. Kallas, E. Hurt, The NUG1 GTPase reveals and N-terminal RNA-binding domain that is essential for association with 60 S pre-ribosomal particles. *J. Biol. Chem.* **281**, 24737–24744 (2006). doi: 10.1074/jbc.M604261200; pmid: 16803892
40. C. Barrio-Garcia *et al*., Architecture of the Rix1-Rea1 checkpoint machinery during pre-60S-ribosome remodeling. *Nat. Struct. Mol. Biol.* **23**, 37–44 (2016). doi: 10.1038/nsmb.3132; pmid: 26619264
41. Z. Chen *et al*., Structural insights into Mdn1, an essential AAA protein required for ribosome biogenesis. *Cell* **175**, 822–834. e18 (2018). doi: 10.1016/j.cell.2018.09.015; pmid: 30318141
42. P. Sosnowski *et al*., The CryoEM structure of the *Saccharomyces cerevisiae* ribosome maturation factor Rea1. *eLife* **7**, e39163 (2018). doi: 10.7554/eLife.39163; pmid: 30460895
43. M. Thoms *et al*., Suppressor mutations in Rpf2-Rrs1 or Rpl5 bypass the Cgr1 function for pre-ribosomal 5S RNP-rotation. *Nat. Commun.* **9**, 4094 (2018). doi: 10.1038/s41467-018-06660-w; pmid: 30291245
44. F. F. M. Bagatelli, F. N. de Luna Vitorino, J. P. C. da Cunha, C. C. Oliveira, The ribosome assembly factor Nop53 has a structural role in the formation of nuclear pre-60S intermediates, affecting late maturation events. *Nucleic Acids Res.* **49**, 7053–7074 (2021). doi: 10.1093/nar/gkab494; pmid: 34125911
45. M. Thoms *et al*., The exosome is recruited to RNA substrates through specific adaptor proteins. *Cell* **162**, 1029–1038 (2015). doi: 10.1016/j.cell.2015.07.060; pmid: 26317469
46. S. Falk *et al*., Structural insights into the interaction of the nuclear exosome helicase Mtr4 with the preribosomal protein Nop53. *RNA* **23**, 1780–1787 (2017). doi: 10.1261/rna.062901.117; pmid: 28883156
47. S. Miyao, K. Saito, R. Oshima, K. Kawahara, M. Nagahama, MTR4 adaptor PICT1 functions in two distinct steps during pre-rRNA processing. *Biochem. Biophys. Res. Commun.* **637**, 203–209 (2022). doi: 10.1016/j.bbrc.2022.11.018; pmid: 36403484
48. J. Gordon *et al*., Cryo-EM reveals the architecture of the PELP1-WDR18 molecular scaffold. *Nat. Commun.* **13**, 6783 (2022). doi: 10.1038/s41467-022-34610-0; pmid: 36351913
49. N. Raman, E. Weir, S. Müller, The AAA ATPase MDN1 acts as a SUMO-targeted regulator in mammalian pre-ribosome remodeling. *Mol. Cell* **64**, 607–615 (2016). doi: 10.1016/j.molcel.2016.09.039; pmid: 27814492
50. S.-T. Mullineux, D. L. J. Lafontaine, Mapping the cleavage sites on mammalian pre-rRNAs: Where do we stand? *Biochimie* **94**, 1521–1532 (2012). doi: 10.1016/j.biochi.2012.02.001; pmid: 22342225
51. S. Singh, A. Vanden Broeck, L. Miller, M. Chaker-Margot, S. Klinge, Nucleolar maturation of the human small subunit processome. *Science* **373**, eabj5338 (2021). doi: 10.1126/science.abj5338; pmid: 34516797
52. B. Michot, N. Joseph, S. Mazan, J. P. Bachellerie, Evolutionarily conserved structural features in the ITS2 of mammalian pre-rRNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. *Nucleic Acids Res.* **27**, 2271–2282 (1999). doi: 10.1093/nar/27.11.2271; pmid: 10325414
53. O. Rodríguez-Galán, J. J. García-Gómez, D. Kressler, J. de la Cruz, Immature large ribosomal subunits containing the 7S pre-rRNA can engage in translation in Saccharomyces cerevisiae. *RNA Biol.* **12**, 838–846 (2015). doi: 10.1080/15476286.2015.1058477; pmid: 26151772
54. A. Sarkar *et al*., Preribosomes escaping from the nucleus are caught during translation by cytoplasmic quality control. *Nat. Struct. Mol. Biol.* **24**, 1107–1115 (2017). doi: 10.1038/nsmb.3495; pmid: 29083413
55. S. Biedka *et al*., Hierarchical recruitment of ribosomal proteins and assembly factors remodels nucleolar pre-60S ribosomes. *J. Cell Biol.* **217**, 2503–2518 (2018). doi: 10.1083/jcb.201711037; pmid: 29691304
56. M. Taoka *et al*., Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res.* **46**, 9289–9298 (2018). doi: 10.1093/nar/gky811; pmid: 30202881
57. A. Faille, K. C. Dent, S. Pellegrino, P. Jaako, A. J. Warren, bioRxiv, in press, doi: 10.1101/2023.02.28.530191.
58. T. Waku *et al*., NML-mediated rRNA base methylation links ribosomal subunit formation to cell proliferation in a p53-dependent manner. *J. Cell Sci.* **129**, 2382–2393 (2016). doi: 10.1242/jcs.183723; pmid: 27149924
59. S. Sharma *et al*., Identification of novel methyltransferases, Bmt5 and Bmt6, responsible for the m3U methylations of 25S rRNA in Saccharomyces cerevisiae. *Nucleic Acids Res.* **42**, 3246–3260 (2014). doi: 10.1093/nar/gkt1281; pmid: 24335083
60. H. Ma *et al*., $N^{6}$-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. *Nat. Chem. Biol.* **15**, 88–94 (2019). doi: 10.1038/s41589-018-0184-3; pmid: 30531910
61. H. Liao *et al*., Human NOP2/NSUN1 regulates ribosome biogenesis through non-catalytic complex formation with box C/D snoRNPs. *Nucleic Acids Res.* **50**, 10695–10716 (2022). doi: 10.1093/nar/gkac817; pmid: 36161484
62. K. E. Sloan *et al*., Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* **14**, 1138–1152 (2017). doi: 10.1080/15476286.2016.1259781; pmid: 27911188
63. J. Georgeson, S. Schwartz, The ribosome epitranscriptome: Inert-or a platform for functional plasticity? *RNA* **27**, 1293–1301 (2021). doi: 10.1261/rna.078859.121; pmid: 34312287
64. J. C. Ulirsch *et al*., The Genetic landscape of Diamond-Blackfan anemia. *Am. J. Hum. Genet.* **104**, 356 (2019). doi: 10.1016/j.ajhg.2018.12.011; pmid: 30735661
65. A. Aspesi, S. R. Ellis, Rare ribosomopathies: Insights into mechanisms of cancer. *Nat. Rev. Cancer* **19**, 228–238 (2019). doi: 10.1038/s41568-019-0105-0; pmid: 30670820
66. K. R. Kampen, S. O. Sulima, S. Vereecke, K. D. Keersmaecker, Hallmarks of ribosomopathies. *Nucleic Acids Res.* **48**, 1013–1028 (2020). doi: 10.1093/nar/gkz637; pmid: 31350888
67. L. Fromm *et al*., Reconstitution of the complete pathway of ITS2 processing at the pre-ribosome. *Nat. Commun.* **8**, 1787 (2017). doi: 10.1038/s41467-017-01786-9; pmid: 29176610
68. Y. Zhao, J. Rai, H. Yu, H. Li, CryoEM structures of pseudouridine-free ribosome suggest impacts of chemical modifications on ribosome conformations. *Structure* **30**, 983–992.e5 (2022). doi: 10.1016/j.str.2022.04.002; pmid: 35489333
69. Y. Motorin, M. Quinternet, W. Rhalloussi, V. Marchand, Constitutive and variable 2′-O-methylation (Nm) in human ribosomal RNA. *RNA Biol.* **18** (sup1), 88–97 (2021). doi: 10.1080/15476286.2021.1974750; pmid: 34503375

## ACKNOWLEDGMENTS

We thank M. Ebrahim, J. Sotiris, and H. Ng for help with grid screening and data collection at the Evelyn Gruss Lipper Cryo-Electron Microscopy Resource Center, The Rockefeller University; S. Heissel for swift mass spectrometry analysis of our purified samples at the Proteomics Resource Center, The Rockefeller University; and members of the Klinge laboratory for critical reading of this manuscript. Molecular graphics and analyses were performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health grant R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. **Funding:** A.V.B. is supported by an EMBO long-term fellowship (ALTF 711-2019) and by funds from a Pels Family Center Postdoctoral Fellowship at The Rockefeller University. S.K. is supported by the G. Harold and Leila Y. Mathers Foundation (grant MF-2104-01554). **Author contributions:** A.V.B. and S.K. conceived of the study, designed the experiments, and analyzed the data. A.V.B. performed the genome editing, prepared the samples of human pre-60S, and performed the cryo-EM analyses. A.V.B. built the atomic models with help from S.K. A.V.B. engineered the rDNA plasmids and performed the functional assays. A.V.B. and S.K. wrote and edited the manuscript. **Competing interests:** The Rockefeller University has filed a patent related to the described human genome editing platform, on which S.K. is listed as an inventor. **Data and materials availability:** The raw unaligned multiframe movies and the consensus pre-60S polished particles stack have been deposited in the Electron Microscopy Public Image Archive (EMPIAR-11472). The composite cryo-EM maps and atomic models have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB), respectively: state A1 (EMD-29252 and 8FKP), state A2 (EMD-29253 and 8FKQ), state B1 (EMD-29254 and 8FKR), state B2 (EMD-29255 and 8FKS), state C1 (EMD-29256 and 8FKT), state C2 (EMD-29257 and 8FKU), state D1 (EMD-29258 and 8FKV), state D2 (EMD-29259 and 8FKW), state E (EMD-29260 and 8FKX), state F (EMD-29261 and 8FKY), state G (EMD-29262 and 8FKZ), state H (EMD-29263 and 8FL0), state I1 (EMD-29265 and 8FL2), state I2 (EMD-29266 and 8FL3), state I3 (EMD-29267 and 8FL4), state J1 (EMD-29268 and 8FL6), state J2 (EMD-29269 and 8FL7), state J3 (EMD-29271 and 8FL9), state K1 (EMD-29272 and 8FLA), state K2 (EMD-29273 and 8FLB), state K3 (EMD-29274 and 8FLC), state L1 (EMD-29275 and 8FLD), state L2 (EMD-29276 and 8FLE), state L3 (EMD-29277 and 8FLF). All consensus and focused refinement maps have also been deposited in the EMDB (tables S1 and S2). Materials in the main text or the supplement are available from S.K. upon reasonable request. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh3892
Materials and Methods
Supplementary Text
Figs. S1 to S37
Tables S1 to S5
References (*70*–*100*)
Data S1 to S4

View/request a protocol for this paper from *Bio-protocol*.

Submitted 1 March 2023; resubmitted 13 March 2023
Accepted 20 May 2023
10.1126/science.adh3892

RESEARCH ARTICLE

PHYSICS

# An improved bound on the electron's electric dipole moment

Tanya S. Roussy[1,2]†, Luke Caldwell[1,2]†, Trevor Wright[1,2], William B. Cairncross[1,2]‡, Yuval Shagam[1,2]§, Kia Boon Ng[1,2], Noah Schlossberger[1,2], Sun Yool Park[1,2], Anzhou Wang[1,2], Jun Ye[1,2], Eric A. Cornell[1,2]*

**The imbalance of matter and antimatter in our Universe provides compelling motivation to search for undiscovered particles that violate charge-parity symmetry. Interactions with vacuum fluctuations of the fields associated with these new particles will induce an electric dipole moment of the electron (eEDM). We present the most precise measurement yet of the eEDM using electrons confined inside molecular ions, subjected to a huge intramolecular electric field, and evolving coherently for up to 3 seconds. Our result is consistent with zero and improves on the previous best upper bound by a factor of ~2.4. Our results provide constraints on broad classes of new physics above $10^{13}$ electron volts, beyond the direct reach of the current particle colliders or those likely to be available in the coming decades.**

Electric dipole moments of fundamental particles, such as the electron, are signatures of time-reversal symmetry violation, equivalent to violation of combined charge and parity (CP) symmetry (*1*). CP symmetry is broken in the standard model but only in the quark sector (*2*), so the coupling to leptons is weak and the predicted electron's electric dipole moment (eEDM) several orders of magnitude below current experimental sensitivity (*3*, *4*). Explaining the imbalance of matter and antimatter in the Universe requires additional CP violation, beyond that present in the Standard Model (*5*–*7*). Many proposed extensions predict new particles at energies higher than any so far discovered, with CP-violating interactions. These new particles can induce a much larger eEDM, often within reach of near-term experiments (*8*–*10*). A nonzero measurement at current experimental sensitivities would unambiguously signal new physics, whereas a more precise measurement consistent with zero imposes challenging constraints on possible explanations of the matter-antimatter imbalance. Our measurement uses quantum projection–noise-limited spectroscopy on samples of hundreds of molecular ions with interrogation times of up to 3 s. Our result, $d_e = (-1.3 \pm 2.0_{\mathrm{stat}} \pm 0.6_{\mathrm{sy}})$ $10^{-30}$ $e$ cm, is consistent with zero and gives an upper bound of $|d_e| < 4.1 \times 10^{-30}$ $e$ cm at 90% confidence.

An eEDM $\vec{d}_e = d_e\hat{s}$—with $\hat{s}$ a unit vector along the spin of the electron—subject to an electric field, $\vec{\mathcal{E}}$, has an energy of $-\vec{d}_e \cdot \vec{\mathcal{E}}$. The essence of an eEDM search is to measure the energy shift when $\hat{s}$ is aligned with $\vec{\mathcal{E}}$ compared with when it is antialigned. The size of the observable shift scales with the size of $\vec{\mathcal{E}}$, and thus many existing (*11*–*13*) and proposed (*14*–*17*) eEDM experiments use electrons embedded inside polar molecules, where intramolecular electric fields can be ~$10^5$ times larger than what can be directly applied in the lab. These internal electric fields can be aligned in the lab frame by orienting the molecules with modest external electric fields.

Our measurement uses $HfF^+$ molecular ions. In an applied electric field of ~58 V cm$^{-1}$, the $^3\Delta_1$ ($v = 0$, $J = 1$) "science" state of the molecule is split into a series of doublets (Fig. 1A). In two of these doublets, highlighted in color, the molecule is oriented (*18*); the upper doublet (orange) has the intramolecular axis parallel to the applied field, whereas the lower doublet (blue) is antiparallel. This intramolecular axis defines the direction of an effective

[1]JILA, NIST and University of Colorado, Boulder, CO 80309, USA. [2]Department of Physics, University of Colorado, Boulder, CO 80309, USA.
†These authors contributed equally to this work.
‡Present address: Atom Computing, Berkeley, CA 94710, USA.
§Present address: Schulich Faculty of Chemistry, Technion–Israel Institute of Technology, Haifa 3200003, Israel.
*Corresponding author. Email: cornell@jila.colorado.edu

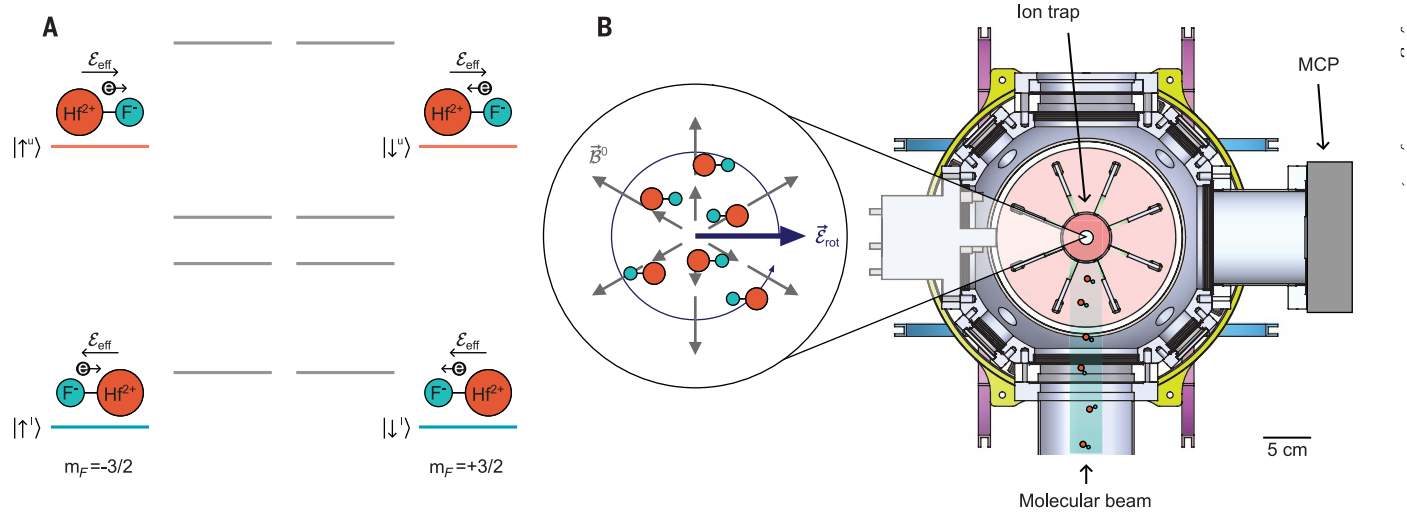



**Fig. 1. Experiment outline.** (**A**) Level structure of the eEDM-sensitive $^3\Delta_1$ ($v = 0, J = 1$) state. The horizontal axis indicates $m_F$, the projection of the total angular momentum onto the externally applied electric field. The vertical axis indicates the energy of the states. The direction of the electron spin and effective electric field, $\mathcal{E}_{\mathrm{eff}}$, is indicated for each of the states used in the experiment. (**B**) Schematic of ion trap, composed of eight radial electrodes and a pair of endcap electrodes. (Inset) Fields applied during experimental sequence: the rotating electric bias field, $\vec{\mathcal{E}}_{\mathrm{rot}}$, and the quadrupole magnetic field, $\vec{\mathcal{B}}^0$.

electric field, $\mathcal{E}_{\mathrm{eff}} \approx 23$ GV cm$^{-1}$ (*19*–*22*), acting on the spin of one of the valence electrons. In the presence of a small magnetic field, the two states in a doublet correspond to the spin of this valence electron being aligned or anti-aligned with $\mathcal{E}_{\mathrm{eff}}$. We prepare a coherent superposition of the two spin states and measure the energy difference using Ramsey spectroscopy. The eEDM will give a contribution to this energy, $\pm 2 d_e \mathcal{E}_{\mathrm{eff}}$, with opposite sign in the two doublets. We perform the measurement simultaneously on spatially overlapping clouds of ions prepared in each of the doublets. The difference between the measured energies is our science signal.

### Experimental overview

Our experimental apparatus is shown schematically in Fig. 1B. An overview of the experimental sequence is given here; more details, including an account of improvements made since our earlier result (*12*), are presented in (*23*, *24*). The sequence begins with production and radiofrequency trapping of roughly 20,000 $HfF^+$ ions. To orient the molecules while maintaining confinement, we rotate the orienting field, $\vec{\mathcal{E}}_{\mathrm{rot}}$, at angular frequency $\omega_{\mathrm{rot}} = 2\pi \times 375$ kHz and perform our spectroscopy in this rotating frame. We also apply a quadrupole magnetic field gradient, $\vec{\mathcal{B}}^0$, to create a time-averaged effective bias magnetic field, $\mathcal{B}_{\mathrm{rot}}$ (*24*).

We prepare an incoherent mixture of one of the spin states from each doublet, either $|\uparrow^{\mathrm{u}}\rangle$ and $|\uparrow^{\mathrm{l}}\rangle$ or $|\downarrow^{\mathrm{u}}\rangle$ and $|\downarrow^{\mathrm{l}}\rangle$ (Fig. 1A), and then apply a $\frac{\pi}{2}$ pulse to create a coherent superposition of the two states in each doublet. We allow the superpositions to evolve for a variable amount of time, then apply a second $\frac{\pi}{2}$ pulse to map the accumulated relative phase between the states in a doublet onto a population difference between those states. We clean out the population in one of the spin states in each doublet, either $|\uparrow^{\mathrm{u/l}}\rangle$ or $|\downarrow^{\mathrm{u/l}}\rangle$, then count the number of ions in the remaining stretched states by state-selectively photodissociating the molecules and detecting the resultant $Hf^+$ ions (*25*). We use the opposing orientations of the two doublets in the trap to send the $Hf^+$ ions originating from molecules in each doublet to opposite sides of our imaging microchannel plate (MCP) and phosphor screen assembly (*26*, *27*). We then repeat the procedure with the opposite initial

**Table 1. Summary of systematic shifts and their uncertainties.** Data are as presented in (*23*). All values are in microhertz.

| Effect | Correction | Uncertainty |
|---|---|---|
| Magnetic | | |
| Nonreversing $\vec{\mathcal{B}}^0$ | 0.1 | $< 0.1$ |
| Stray $\mathcal{B}$ fields + distortion of $\mathcal{E}_{\mathrm{rot}}$ | | 3.2 |
| Berry's phase | | |
| Rotation-odd axial secular motion | | 3.4 |
| Axial fields at harmonics of $\mathcal{E}_{\mathrm{rot}}$ | | 3.4 |
| Simultaneous doublet spectroscopy | | |
| Imperfect spatial overlap | | 3.5 |
| Imperfect imaging contrast | | 1.4 |
| Other | | |
| Rotation-induced $m_F$-level mixing | | 0.4 |
| Total | 0.1 | 6.9 |

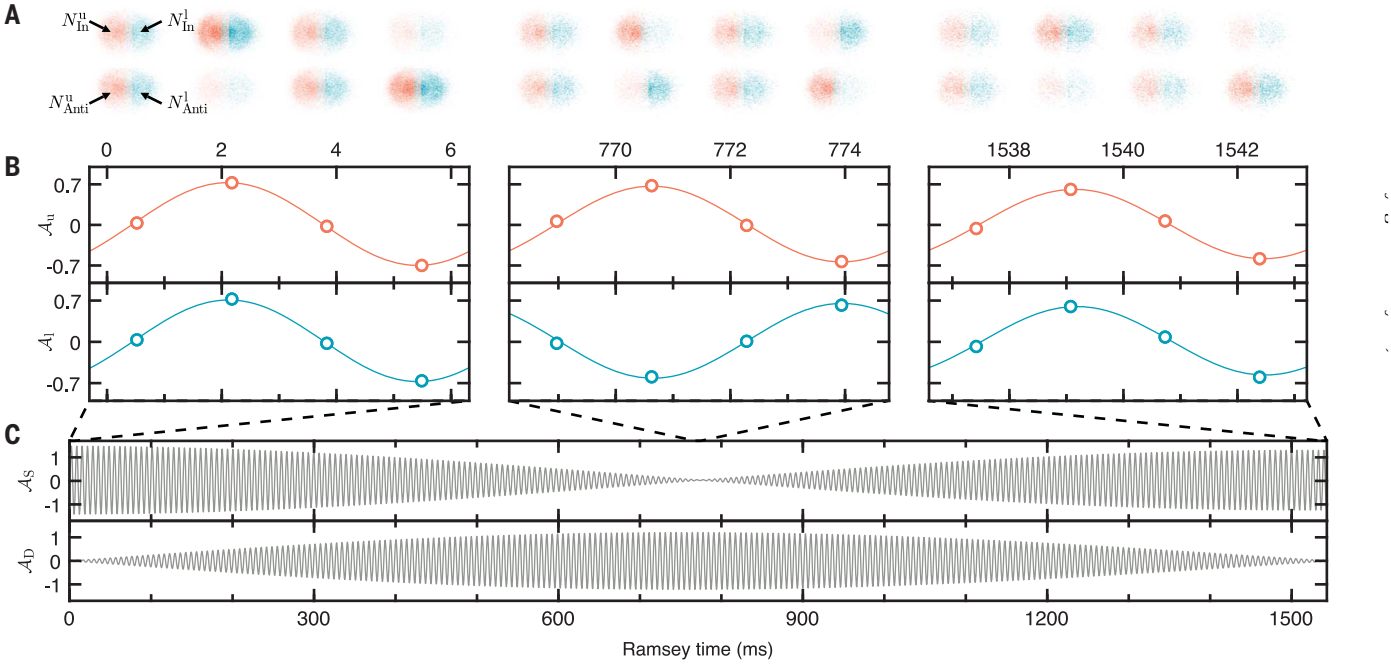



**Fig. 2. Example Ramsey data.** (**A**) Detection of $Hf^+$ ions; ions are assigned to the upper or lower doublet based on their position (upper doublet shown in orange, lower doublet in blue). Counts from a thin central swatch where the assignment is ambiguous (shown in gray) are removed. Images shown are averaged over 60 shots of the experiment. (**B**) Asymmetries for the upper and lower doublet. (**C**) Fitted sum and difference asymmetries, $\mathcal{A}_{\mathrm{S}}$ and $\mathcal{A}_{\mathrm{D}}$, used to extract mean and difference frequencies, $f_{\mathrm{m}}$ and $f_{\mathrm{d}}$. Middle-time data, where the two doublets are out of phase, were collected in this example dataset for illustrative purposes only. Such data contribute very little to the frequency determination and were not collected during the final precision dataset.

spin state. Example data are shown in Fig. 2A. We count the ions on each side of the screen, in each configuration—typically ~120 ions from each doublet after a 3-s hold time—and from these two measurements construct the two asymmetries, $\mathcal{A}_u$ and $\mathcal{A}_l$, where

$$\mathcal{A}_{u/l} = \frac{N_{In}^{u/l} - N_{Anti}^{u/l}}{N_{In}^{u/l} + N_{Anti}^{u/l}} \quad (1)$$

Here, $N_{In}^{u/l}$ and $N_{Anti}^{u/l}$ are the number of ions counted, u/l indicates the upper or lower doublet, and the subscripts indicate whether we read out the same state that we prepare (In) or the opposite (Anti). We repeat our measurement at different free-evolution times, generating a pair of Ramsey fringes as shown in Fig. 2B. The frequencies of these two fringes are proportional to the energy splitting in the two doublets. The primary contribution to this energy splitting is the Zeeman splitting, $3g_F\mu_B\mathcal{B}_{rot}$, where $g_F = -0.0031(1)$ is the $g$ factor of the science state (*28*). For our typical experimental parameters, this produces fringe frequencies of ~100Hz. Other effects, including the eEDM, make small modifications to this frequency.

Instability of the intensity of the pulsed lasers used for creation and photodissociation of the ions (*24*) creates considerable noise in the number of ions measured in each shot of the experiment, typically approximately three times as large as the quantum projection–noise limit on the side of the fringe at 3 s. However, these sources of noise, and many others, are common mode between the two doublets—the exact same laser pulses address both clouds of ions—and so the noise in $\mathcal{A}_u$ and $\mathcal{A}_l$ is highly correlated. To take advantage of this, we form the sum and difference asymmetries $\mathcal{A}_S = \mathcal{A}_u + \mathcal{A}_l$ and $\mathcal{A}_D = \mathcal{A}_u - \mathcal{A}_l$ (Fig. 2C). If we take data when the two doublets are close to being in phase, the noise in $\mathcal{A}_D$ is drastically reduced (*27*). The two doublets oscillate at slightly different frequencies, $f_u$ and $f_l$, owing to a ~1/230 fractional difference in their magnetic moments, and so during the eEDM dataset we deliberately take our data at a beat. We take two sets of points: the early-time data, when the two doublets are in phase, and the late-time data ~230 oscillations later, when they come back into phase again. We can control the time of the second beat by varying the strength of the magnetic bias field, $\mathcal{B}_{rot}$. We fit to $\mathcal{A}_S$ and $\mathcal{A}_D$ to extract the mean of the two fringe frequencies, $f_m = \frac{1}{2}(f_u + f_l)$, and their difference, $f_d = \frac{1}{2}(f_u - f_l)$.

We collect Ramsey fringes in $2^3 = 8$ experimental states, corresponding to each possible combination of three binary experimental switches, $\{\tilde{B}, \tilde{R}, \tilde{I}\} = \pm 1$. $\tilde{B}$ is the direction of the magnetic bias field relative to $\vec{\mathcal{E}}_{rot}$, $\tilde{R}$ the rotation direction of $\vec{\mathcal{E}}_{rot}$, and $\tilde{I}$ the direction of $\vec{\mathcal{E}}_{rot}$ relative to the imaging MCP at the instant of photodissociation, determining which side of the phosphor screen each of the doublets is imaged onto. A set of Ramsey fringes in each of the 8 switch states forms a block. To minimize the effects of experimental drifts, within a block we interleave data collection for the switch states; the first Ramsey time is recorded for all switch states before moving onto the second Ramsey time for each switch state, and so on. We take the 16 fitted frequencies from each block and form 16 linear combinations to give the components of the measured frequencies, which are even or odd under each of the experimental switches. Following (*29*), we label the components with superscripts that denote the switches under which the quantity is odd. For example, our science signal is $f^{DB}$, the component of the difference frequency that is odd under $\tilde{B}$ but even under $\tilde{R}$ and $\tilde{I}$ (*30*). The other channels allow us to diagnose systematics and monitor experimental performance.

Over the course of the dataset, we varied a number of other experimental parameters on timescales slower than a block. These include the state we read out at the end of the Ramsey sequence, denoted $\tilde{P} = \pm 1$ and alternated each block; the order in which the switch states are recorded at each Ramsey time, alternated every other block; three different magnitudes of the magnetic bias field, corresponding to mean fringe frequencies of $f^0$~77, 105, and 151 Hz; and reversal of the waveplates that set the lab-frame handedness of the light used for state preparation and readout. During data collection and analysis, we "blinded" our measurement of $f^{DB}$ by adding an unknown offset to this channel. The offset was not removed until our systematics search and analysis (*23*) were complete.

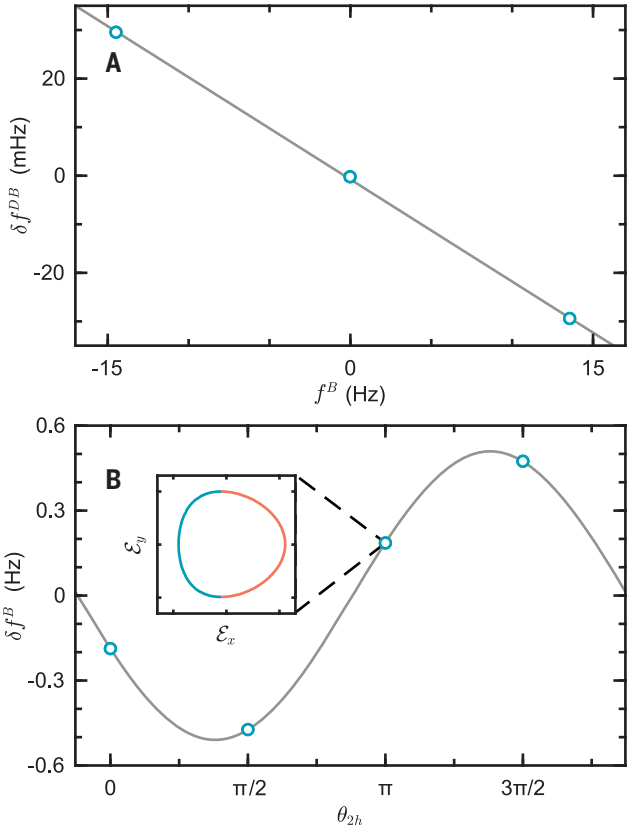



**Fig. 3. Systematic shifts in the measurement.** (**A**) Shift in $f^{DB}$ caused by nonreversing magnetic quadrupole field, which can be corrected by using the shift in the $f^B$ channel, which is ~460 times as large as the $f^{DB}$shift. (**B**) Shift in the $f^B$ channel caused by deliberately applied second-harmonic electric field $\mathcal{E}_{2h}$ with transverse magnetic field, $\mathcal{B}$. Data show variation in shift as angle $\theta_{2h}$ between $\mathcal{E}_{2h}$ and $x$ axis is varied. $\mathcal{E}_{2h}$ is ~250 times as large as that present in the experiment; $\mathcal{B}$ is ~ 14 mG. (Inset) Lissajous figure traced out by total electric field for greatly exaggerated ratio of $\mathcal{E}_{2h}/\mathcal{E}_{rot}$; blue and orange show one-half cycle each. The field points in the $-x$ direction for more time than in the $+x$ direction.

## Accuracy evaluation

To evaluate the accuracy of our measurement, we searched extensively for systematic shifts before data collection; a summary is given in Table 1. In general, we tuned a variety of experimental parameters over ranges that were large compared with those present during data collection, exaggerating any accompanying systematic effects, and observed the response in our data channels. The only shift we could observe directly in the eEDM channel stems from a nonreversing quadrupole magnetic field and the difference in magnetic moments between the two doublets, caused primarily by the applied electric field mixing the states of the two doublets with higher rotational levels of the molecule. The $f^B$ channel provides a direct measurement of the nonreversing magnetic field and allows us to apply a correction to our science channel, $\delta f_{corr}^{DB} = f^B \frac{\delta g_F}{g_F}$, where $\delta g_F$ is half the difference between the $g$ factors for the upper and lower doublets (Fig. 3A). Before applying any corrections to the science channel, we suppress this systematic by actively shimming the currents through

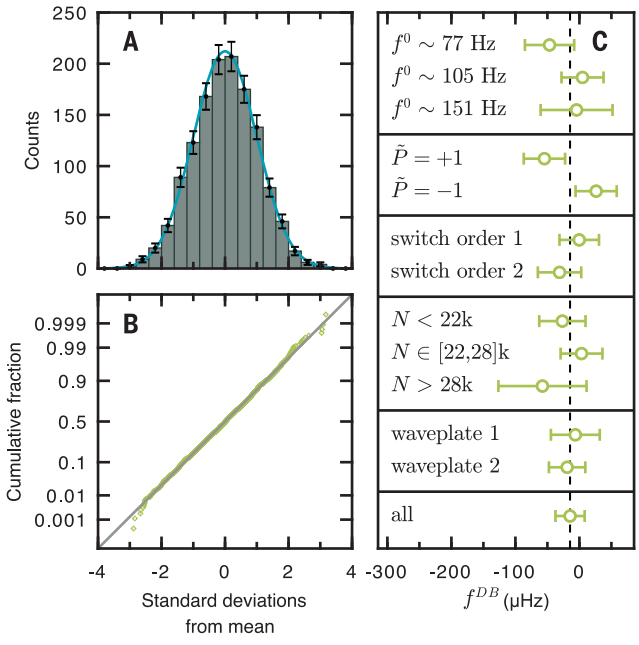



**Fig. 4. Summary of our dataset.** Cuts have been applied and the uncertainty on each block scaled by $\sqrt{\chi^2}$ to account for overscatter. (**A**) Histogram of data. Error bars show standard deviation of bin counts expected from Poisson distribution. The blue line shows normal distribution. (**B**) Normal probability plot of $f^{DB}$, showing that the data are consistent with a normal distribution. The gray line shows expected probability for a normal distribution. (**C**) Variation of central value under different experimental parameters compared with the overall average value of $f^{DB}$. Here, $N$ is the average number of trapped $HfF^+$ ions per experimental trial during a block. Other panels in (C) are described in the final paragraph of the Experimental Overview section.

the coils that apply the magnetic bias field to minimize $f^B$. This shimming was so effective that the mean correction that we applied was well below our statistical sensitivity, $\delta f^{DB}_{\mathrm{corr}} = 90$ nHz.

The shimming and correction procedure leaves us susceptible to other possible effects that cause shifts in $f^B$ and $f^{DB}$ with a ratio different from $\frac{\delta g_F}{g_F}$. An important example of such a shift is the combination of a transverse magnetic field with an electric field oscillating at $2\omega_{\mathrm{rot}}$, which was present in our experiment owing to harmonic distortion in the amplifiers driving the trap electrodes. The harmonic distortion causes the electric field, and thus the magnetic moment of the molecule, to spend more time pointing in one spatial direction than the other, giving a nonzero time-averaged Zeeman interaction with a background magnetic field. This causes shifts in $f^B$ but no corresponding shifts in $f^{DB}$, where the effect is canceled by the coincident change in the size of the differential magnetic moment owing to the distortion. Figure 3B shows the shift in $f^B$ when we deliberately apply a second-harmonic electric field and vary its angle. We shim out the second harmonic on each electrode by feeding forward a second-harmonic signal with the opposite phase, suppressing the amplitude by ~80 dB. We used magnetic shim coils to null the ambient magnetic field at the trap center to <10 mG. The measured sizes of the residual effects were used to compute the maximum size of the systematic during our dataset.

A full account of all systematic shifts considered is presented in (*23*).

### Measuring the eEDM

We collected 1370 blocks over about 2 months, corresponding to ~620 hours of data and ~$10^8$ ion detection events. Each block results in one value of $f^{DB}$ and thus a single measurement of $d_e$. The uncertainty on $f^{DB}$ for each block is calculated with only the standard errors on the asymmetries for that block. We applied cuts to the blinded data on the basis of non-eEDM channels that indicated signal quality. Blocks with late-time contrast <0.2 were cut because of a low signal-to-noise ratio, as were blocks containing fitted fringe frequencies that were >3.5σ different from the mean fringe frequency for that switch state. After applying cuts, we were left with 1329 blocks with $\chi^2 = 1.07(4)$ for $f^{DB}$. Figure 4, A and B, show the distribution of measured $f^{DB}$ values over the 1329 blocks after relaxing the uncertainty for each of the blocks by a factor of $\sqrt{\chi^2} = 1.035$. The data are consistent with a normal distribution. Our final statistical uncertainty of 22.8 μHz is obtained with these relaxed uncertainties. Based on the number of ions detected in each shot, this uncertainty is ~30% above the quantum projection–noise limit. More details on how we determine uncertainties are given in (*24*).

Figure 4C shows how the measured value of $f^{DB}$ depends on experimental parameters varied during the dataset; we find no concerning dependencies.

We removed our blind on 1 November 2022, and obtained a final value for the eEDM-sensitive frequency channel

$$f^{DB} = -14.6 \pm 22.8_{\mathrm{stat}} \pm 6.9_{\mathrm{syst}}\ \mu\mathrm{Hz} \quad (2)$$

Dividing by $-2\mathcal{E}_{\mathrm{eff}}\frac{\mathrm{sgn}(g_F)}{h} \simeq 1.11 \times 10^{31}\ \mu\mathrm{Hz}\ e^{-1}\,\mathrm{cm}^{-1}$ (*21*, *31*), we obtain a value for the eEDM

$$d_e = (-1.3 \pm 2.0_{\mathrm{stat}} \pm 0.6_{\mathrm{syst}}) \times 10^{-30}\ e\ \mathrm{cm} \quad (3)$$

which is consistent with zero within one standard error. The combined statistical and systematic uncertainty, $\sigma_{d_e} = 2.1 \times 10^{-30}\ e$ cm, improves on our previous work (*12*) by a factor of ~37, and on the previous state-of-the-art from the ACME collaboration (*13*) by a factor of ~2. This result and that of the ACME collaboration—two measurements using very different experimental platforms with contrasting sources of systematic shifts—are consistent at slightly above one standard error.

### Discussion

We use our result to obtain an upper bound using a folded Gaussian distribution

$$|d_e| < 4.1 \times 10^{-30}\ e\ \mathrm{cm}\ \ (90\%\ \ \mathrm{confidence}) \quad (4)$$

This limit constrains extensions to the Standard Model that predict new sources of CP-symmetry violation to explain the matter-antimatter asymmetry of the Universe (*32*). Many extensions, including supersymmetry, the two-Higgs model, and left-right symmetric models, generate an eEDM at the one-loop level (*33*), with magnitude (*9*)

$$d_e \sim \frac{e a_0 \alpha}{2}\frac{g^2}{2\pi}\sin\phi_{\mathrm{CP}}\frac{m_e^2}{M^2} \quad (5)$$

Here, $M$ is the characteristic mass of new particles with effective coupling strength, $g$, to the electron; $\phi_{\mathrm{CP}}$ is the phase that describes how strongly the interaction violates CP symmetry; $m_e$ and $e$ are the mass and charge of the electron respectively; and $\alpha$ is the fine-structure constant. Because $d_e \propto M^{-2}$, and because our limit in Eq. 4 is a factor of ~2.4 smaller than the limit reported in (*13*), we are sensitive to new particles with mass that is $\sqrt{2.4} = 1.5$ times as large.

To estimate the mass reach of our experiment, we need to make assumptions for the size of $g^2$ and $\sin\phi_{\mathrm{CP}}$. For the strong force, quantum electrodynamics, and the weak force, $g^2 \approx 1, 1/137$, and $10^{-6}$, respectively. For extensions to the Standard Model seeking to explain the matter-antimatter asymmetry, the naive expectation is that $\sin\phi_{\mathrm{CP}} \sim 1$. With this assumption, we can interpret our new limit on $d_e$ as $M \gtrsim \left(g/\alpha^{\frac{1}{2}}\right)$ 40 TeV. For new particles with $g^2$ of order $\alpha \sim 1/137$, this bound is an order of magnitude greater than the largest-mass particles that can be directly detected at the Large Hadron Collider (*34*).

So far, we have assumed that CP violation arises purely from $d_e$. Diatomic molecules are also sensitive to pseudoscalar-scalar electron-nucleon coupling, $C_S$ (*35*, *36*), and we can interpret our measurement as a linear combination, $hf^{DB} \times \mathrm{sgn}(g_F) = -2\mathcal{E}_{\mathrm{eff}} d_e + 2W_S C_S$, where $\frac{W_S}{h} = -51$ kHz (*31*) is a molecule-specific structure constant. Assuming that $d_e$ is zero, we can instead attribute our measurement to $C_S$, and we find

$$C_S = (-1.4 \pm 2.2_{\mathrm{stat}} \pm 0.7_{\mathrm{syst}}) \times 10^{-10} \quad (6)$$

Determining rigorous limits on $d_e$ and $C_S$ requires combining the results of two or more measurements using molecules with different ratios of $\mathcal{E}_{\mathrm{eff}}$ to $W_S$. Figure S1 shows a combined fit to the results of this work and (*13*), giving upper bounds of $|d_e| < 2.1 \times 10^{-29}\, e\,\mathrm{cm}$ and $|C_S| < 1.9 \times 10^{-9}$ with 90% confidence. Our measurement improves these bounds by factors of 16 and 12, respectively (*37*).

## REFERENCES AND NOTES


1. I. B. Khriplovich, S. K. Lamoreaux, *CP Violation Without Strangeness: Electric Dipole Moments of Particles, Atoms, and Molecules*, Texts and Monographs in Physics (Springer, 1997).
2. C. Patrignani, *Chin. Phys. C* **40**, 100001 (2016).
3. Y. Yamaguchi, N. Yamanaka, *Phys. Rev. Lett.* **125**, 241802 (2020).
4. Y. Ema, T. Gao, M. Pospelov, *Phys. Rev. Lett.* **129**, 231801 (2022).
5. M. Dine, A. Kusenko, *Rev. Mod. Phys.* **76**, 1–30 (2003).
6. A. D. Sakharov, *Pis'ma Z. Eksp. Teor. Fiz.* **5**, 32–35 (1967).
7. M. B. Gavela, P. Hernández, J. Orloff, O. Pène, *Mod. Phys. Lett. A* **09**, 795–809 (1994).
8. M. Pospelov, A. Ritz, *Ann. Phys.* **318**, 119–169 (2005).
9. J. Engel, M. J. Ramsey-Musolf, U. Van Kolck, *Prog. Part. Nucl. Phys.* **71**, 21–74 (2013).
10. Y. Nakai, M. Reece, *J. High Energy Phys.* **2017**, 31 (2017).
11. J. J. Hudson *et al.*, *Nature* **473**, 493–496 (2011).
12. W. B. Cairncross *et al.*, *Phys. Rev. Lett.* **119**, 153001 (2017).
13. ACME Collaboration, *Nature* **562**, 355–360 (2018).
14. I. Kozyryev, N. R. Hutzler, *Phys. Rev. Lett.* **119**, 133002 (2017).
15. A. Vutha, M. Horbatsch, E. Hessels, *Phys. Rev. A* **98**, 032513 (2018).
16. The NL-eEDM collaboration *et al.*, *Eur. Phys. J. D* **72**, 197 (2018).
17. N. Fitch, J. Lim, E. Hinds, B. Sauer, M. Tarbutt, *Quantum Sci. Technol.* **6**, 014006 (2020).
18. A. E. Leanhardt *et al.*, *J. Mol. Spectrosc.* **270**, 1–25 (2011).
19. E. R. Meyer, J. L. Bohn, M. P. Deskevich, *Phys. Rev. A* **73**, 062108 (2006).
20. A. Petrov, N. Mosyagin, T. Isaev, A. Titov, *Phys. Rev. A* **76**, 030501 (2007).
21. L. V. Skripnikov, *J. Chem. Phys.* **147**, 021101 (2017).
22. T. Fleig, *Phys. Rev. A* **96**, 040502 (2017).
23. L. Caldwell *et al.*, Systematic and statistical uncertainty evaluation of the HfF+ electron electric dipole moment experiment. arXiv:2212.11837 [physics.atom-ph] (2022).
24. Materials and methods are available as supplementary materials.
25. K. K. Ni *et al.*, *J. Mol. Spectrosc.* **300**, 12–15 (2014).
26. Y. Shagam *et al.*, *J. Mol. Spectrosc.* **368**, 111257 (2020).
27. Y. Zhou *et al.*, *Phys. Rev. Lett.* **124**, 053201 (2020).
28. H. Loh, "Search for an electron electric dipole moment with trapped molecular ions," thesis, University of Colorado Boulder and JILA (2013).
29. J. Baron *et al.*, *Science* **343**, 269–272 (2014).
30. We define the mean frequency to be always positive. If we had instead allowed the fringe frequency to change sign when the sign of the magnetic bias field changes, our science signal would have been $f^0$.
31. T. Fleig, M. Jung, *J. High Energy Phys.* **2018**, 12 (2018).
32. L. Canetti, M. Drewes, M. Shaposhnikov, *New J. Phys.* **14**, 095012 (2012).
33. T. Chupp, P. Fierlinger, M. Ramsey-Musolf, J. Singh, *Rev. Mod. Phys.* **91**, 015001 (2019).
34. T. Bose *et al.* in *Proceedings of the 2022 Snowmass Summer Study*, Seattle, WA, 17 to 26 July 2022.
35. T. Chupp, M. Ramsey-Musolf, *Phys. Rev. C Nucl. Phys.* **91**, 035502 (2015).
36. M. Denis *et al.*, *New J. Phys.* **17**, 043005 (2015).
37. Improvements are given with respect to equivalent bounds calculated from combined fit to results of (*13*) and (*12*).
38. T. S. Roussy *et al.*, Data presented in "An improved bound on the electron's electric dipole moment," Zenodo (2023); https://doi.org/10.5281/zenodo.7837398.


## ACKNOWLEDGMENTS

We thank the staff at JILA for making this experiment possible. **Funding:** This work was supported by the Cydney and Tom Marsico Family Foundation, the Sloan Foundation, the Gordon and Betty Moore Foundation, NIST, and the NSF (award PHY-1125844). T.W. acknowledges funding support from NSF GRFP. **Author contributions:** All authors contributed, in varying degrees, to the conception of experimental approach, molecular survey spectroscopy, apparatus design and construction, apparatus debugging and commissioning, apparatus maintenance, data collection and analysis, analysis of systematic errors, and preparation of publications. We are all responsible for the accuracy of the bottom-line result and of the written account we have submitted to *Science*. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Data and code used for analysis are available at Zenodo (*38*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg4084
Materials and Methods
Figs. S1 and S2
References (*39*–*42*)

Submitted 22 December 2022; accepted 18 May 2023
10.1126/science.adg4084

BATTERIES

# A lithium superionic conductor for millimeter-thick battery electrode

Yuxiang Li[1]†, Subin Song[2], Hanseul Kim[2], Kuniharu Nomoto[1], Hanvin Kim[1]‡, Xueying Sun[2]§, Satoshi Hori[1], Kota Suzuki[1], Naoki Matsui[1], Masaaki Hirayama[1,2], Teruyasu Mizoguchi[3], Takashi Saito[4,5,6], Takashi Kamiyama[4,5]¶, Ryoji Kanno[1]*

No design rules have yet been established for producing solid electrolytes with a lithium-ion conductivity high enough to replace liquid electrolytes and expand the performance and battery configuration limits of current lithium ion batteries. Taking advantage of the properties of high-entropy materials, we have designed a highly ion-conductive solid electrolyte by increasing the compositional complexity of a known lithium superionic conductor to eliminate ion migration barriers while maintaining the structural framework for superionic conduction. The synthesized phase with a compositional complexity showed an improved ion conductivity. We showed that the highly conductive solid electrolyte enables charge and discharge of a thick lithium-ion battery cathode at room temperature and thus has potential to change conventional battery configurations.

All-solid-state lithium batteries (ASSLBs) have attracted research interest because the solid battery configuration—which uses lithium superionic conductors as the electrolyte instead of liquids in the current lithium batteries—has potential for improved safety and enhanced energy–power characteristics (*1*–*4*). After decades of research all-solid-state cells are capable of discharging at a high current density of >10 mA $cm^{-2}$ (*5*). Although the cathode used in that study was only tens of micrometers thick, such fast discharge properties imply high energy and power densities for future ASSLBs, especially at elevated temperatures such as 100°C in which the operation of Li-ion batteries using organic liquid electrolytes (e.g., those containing $LiPF_6$ and ethylene carbonate) are currently restricted due to safety concerns (*4*, *6*, *7*). These results are attributed to solid electrolytes with $Li^+$ conductivities higher than 10 mS $cm^{-1}$ at room temperature, which is comparable to that of traditional liquid electrolytes (*8*–*10*). These solid electrolytes include $Li_{10}GeP_2S_{12}$ (LGPS) (*8*) and Li-argyrodites represented by $Li_6PS_5Cl$ (*11*). A conductivity of 25 mS $cm^{-1}$ was reported for $Li_{9.54}Si_{1.74}P_{1.44}S_{11.7}Cl_{0.3}$ (LSiPSCl) (*5*) having the same kind of crystal structure as that of LGPS (LGPS-type structure), which is characterized by a *bcc*-like anion sublattice, a framework composed of tetrahedral $[P/Tt]S_4$ units (*Tt* = Si, Ge, Sn), and widely distributed Li sites (*12*, *13*).

Although the conductivities of these solid electrolytes are comparable to those of liquid electrolytes their stiff nature can be a disadvantage. As a result of the difficulty in wetting the surface of the cathode active material as a catholyte, solid electrolytes must be incorporated at the microstructural level into the electrode, resulting in void space inside it. This hinders the homogeneous supply of $Li^+$ to the active material, leading to a capacity loss (*14*) that makes the performance of ASSLBs inferior to that of Li-ion batteries despite the superior conductivity of the solid electrolytes relative to liquid ones. This problem presumably becomes more pronounced in millimeter-thick electrodes, where ohmic drop increases when the conductivity of solid electrolyte is low (*15*). In addition, the capacity decrease of the solid-state battery would be even more severe at low temperatures (below 0°C) in which the conductivity of the electrolyte drops sharply.

The history of the development of inorganic ionic conductors suggests that conductivity enhancement was achieved using the multiple substitution strategy. $Ag^+$ (*16*) and $Cu^+$ (*17*) conductors displayed high conductivities of >200 mS $cm^{-1}$ at room temperature (*17*–*19*) through compositional changes from AgI to $RbAg_4I_5$ or CuI to $Rb_4Cu_{16}I_7Cl_{13}$. This approach corresponds to the "mixed (poly)-anion" (*20*, *21*) or "high-entropy" (*22*, *23*) design reported previously for the development of some $Li^+$ conductors, in which the aim was to flatten the potential barrier for $Li^+$ migration (*21*). The multisubstitution approach is attractive for realizing phases with higher ion conductivity, provided that the materials possess a specific crystal structure having superionic conduction pathways, such as LGPS-type crystal structures. However, there has been no practical principle to realize compositional complexity in the anions and cations or to avoid collapse of the target crystal structure. Consequently, this multi-element substitution strategy has not been used in designing superionic crystals with high conductivity nor have researchers explored the full potential of the crystal structure that enables superionic conduction.

We developed a solid with high $Li^+$ conductivity using a high-entropy material design while preserving the target crystal structure. The simple design rule, as shown in Fig. 1, is derived from the chemical composition trend of well-known argyrodite-type and LGPS-type superionic crystals (*8*, *11*). The diagram shows the plot for two composition parameters: the crystal structure indicator (*t*) and the compositional complexity metric ($S_{mix}$). The former is the ratio of sum of volumes for constituent anions ($V_{sum\text{-}anion}$) to that of cations including lithium ($V_{sum\text{-}cation}$) and can be calculated from

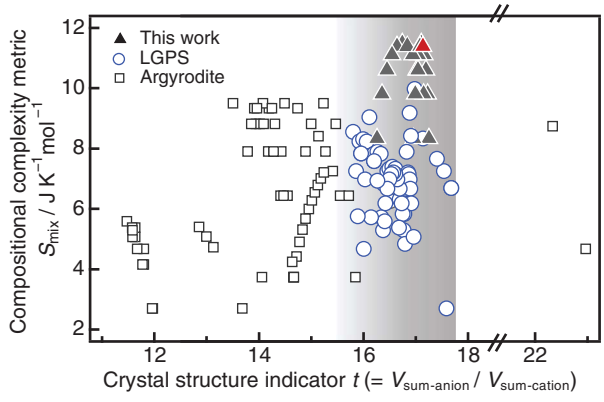



**Fig. 1. Relationship between the crystal structure indicator (*t*, volume ratio of anions to that of cations) and compositional complexity metric ($S_{mix}$) for reported LGPS-type (open circles) and argyrodite-type (open squares) electrolytes.** The newly explored phases in this work are indicated by filled triangles. Gray shading indicates the $V_{sum\text{-}anion}/V_{sum\text{-}cation}$ range for LGPS-type phases and the red filled triangle corresponds to the phase with the highest conductivity.

[1]Research Center for All-Solid-State Battery, Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan. [2]Department of Chemical Science and Engineering, School of Materials and Chemical Technology, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan. [3]Institute of Industrial Science, the University of Tokyo, Tokyo 153-8505, Japan. [4]Neutron Science Division (KENS), Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), 203-1 Shirakata, Tokai, Ibaraki 319-1106, Japan. [5]Department of Materials Structure Science, School of High Energy Accelerator Science, The Graduate University for Advanced Studies, SOKENDAI, 203-1 Shirakata, Tokai, Ibaraki 319-1106, Japan. [6]Japan Proton Accelerator Research Complex (J-PARC) Center, Materials and Life Science Division, 203-1 Shirakata, Tokai, Ibaraki 319-1106, Japan.
*Corresponding author. Email: kanno@echem.titech.ac.jp
†Present address: Advanced Materials and Processing Laboratory, Research Division, Nissan Motor Co., Ltd., 1, Natsushima, Yokosuka, Kanagawa 237-8523, Japan.
‡Present address: Next Generation Battery R&D Center, SK On, 325, Expo-ro, Yuseong-gu, Daejeon, 34124, Republic of Korea.
§Present address: Research department for All-Solid-State Battery, Institute of battery Technology, Mache Power Technology Co., Ltd, 202 Zhendong Road, Huangpu, Guangzhou 510799, P.R. China.
¶Present address: China Spallation Neutron Source (CSNS), Institute of High Energy Physics (IHEP), Chinese Academy of Sciences (CAS) and Spallation Neutron Source Science Center (SNSSC), 1 Zhongziyuan Road, Dalang, Dongguan, Guangdong 523803, P.R. China.

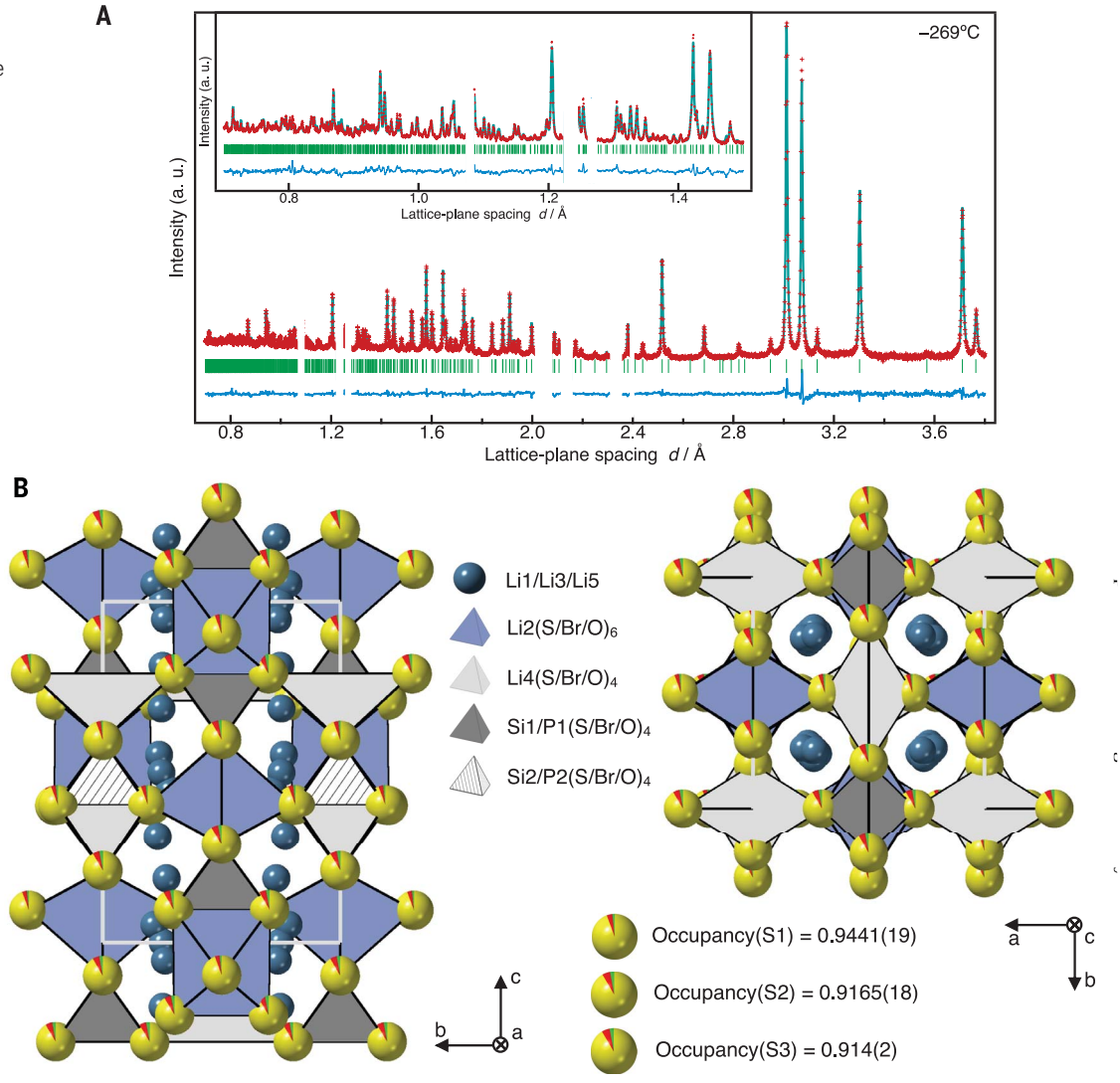



**Fig. 2. Structure analysis of LSiPSBrO.** (**A**) Neutron Rietveld refinement pattern at −269°C. The observed and calculated patterns and the residual differences are shown in red, cyan, and blue, respectively. Green bars indicate the positions of Bragg reflections. Lattice-plane spacing *d* ranges excluded from the analysis are not shown. (**B**) Schematic of the LSiPSBrO crystal structure determined by Rietveld analysis. Occupancy of the S, Br, and O atoms at the S sites are indicated by yellow, green, and red colors, respectively.

the ionic radius (*24*). The latter originates from the configuration entropy of mixing in high-entropy alloys and can be calculated from the occupancy disorder for the anion and cation species (except lithium). The calculation details of $t$ and $S_{mix}$ are presented in figs. S1 to S4. Table S1 lists the $t$ and $S_{mix}$ values for the newly synthesized materials and previously reported argyrodite-type and LGPS-type phases.

The gray shaded region in Fig. 1 indicates that preservation of the LGPS-type structure requires stricter control of the $t$ parameter compared with that for the argyrodite-type structure. Because pure phase construction is the prerequisite for realizing the full potential of compositionally complex materials, we chose a high $S_{mix}$ value and a feasible $t$ value as the design guide for effectively exploring the conductive phases. Based on this rule, in this study we modified the LGPS-type LSiPSCl to design a series of compositions; $Li_{9.54}[Si_{1-\delta}M_{\delta}]_{1.74}P_{1.44}S_{11.1}Br_{0.3}O_{0.6}$ ($LSiM_{\delta}PSBrO$, $M$ = Ge, Sn; δ and 1–δ are the molar ratios of $M^{4+}$ and $Si^{4+}$ in the target composition, respectively, with $0 \leqq \delta \leqq 1$). The designed compositions have high values of $S_{mix}$ (solid triangles in Fig. 1) not achieved by previously reported LGPS- or argyrodite-type material groups (listed in table S1) whereas their $t$ parameters are within the range of previously reported LGPS-type single phases. After systematically synthesizing and characterizing these selected compositions, the highest observed σ value was 32 mS $cm^{-1}$ at 25°C.

## Synthesis and structure analysis

First, we carried out the synthesis and structural analysis for the simplest of the designed $LSiM_{\delta}PSBrO$ series with $\delta = 0$ (i.e., LSiPSBrO). Figure 2A shows the neutron diffraction pattern of LSiPSBrO at −269°C, which was used for structure determination by Rietveld refinement. The pattern was indexed with the same space group as that of LGPS ($P4_2/nmc$) and unchanged at 272°C (fig. S5 and table S2), indicating that the phase maintains an LGPS-type crystal structure up to this high temperature. In the Rietveld refinement the initial structure model was taken from the literature (*12*). Figure 2B shows the crystal structure of LSiPSBrO at −269°C, based on the determined structural parameters listed in table S3. The Br and O atoms in LSiPSBrO prefer to averagely substitute the three sulfur (S) sites. The obtained increase in constituent anion species—which are highly disordered in LSiPSBrO—could potentially lead to energy landscape flattening and a lower activation energy barrier, thus favoring fast ionic migration (*9*, *25*, *26*). This expectation is supported by the energy barrier simulation using ab initio calculation (fig. S6), which indicated that although the substitution degree for S atoms was low (<10 mol%), the influence on the potential energy landscape for $Li^+$ migration could be substantial. Thus, the current high-entropy design may enhance $Li^+$ conduction.

Based on LSiPSBrO, structural analysis was further performed for the Ge- and Sn-substituted $LSiM_{\delta}PSBrO$ samples ($0 < \delta \leqq 1$). As expected from the $t$ parameter, the designed compositions demonstrated reasonably pure LGPS-type phases, with the ratio of more than 85 wt.% as

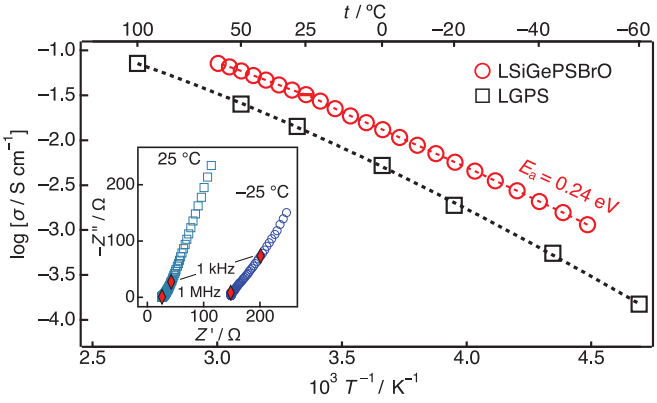



**Fig. 3. Ionic conductivity.** Arrhenius plots for the bulk conductivities of $Li_{9.54}[Si_{0.6}Ge_{0.4}]_{1.74}P_{1.44}S_{11.1}Br_{0.3}O_{0.6}$ and LGPS. Inset: typical impedance plots measured at −25°C (right) and 25°C (left).

determined from Rietveld refinement (table S4). Specifically, LGPS-type single phase was obtained in the range of $0 \leqq \delta \leqq 0.5$ for the Ge-substituted samples and $0 \leqq \delta \leqq 0.4$ for the Sn-substituted samples. The powder x-ray diffraction (XRD) patterns are shown in fig. S7. The lattice parameters were calculated from the synchrotron XRD data and are shown in fig. S8. The unit cell volume increased as the amount of Ge/Sn increased in the LGPS-type LSi$M_\delta$PSBrO phase, confirming the substitution of Si by either Ge or Sn. Thus, the solid-solution limits for these LGPS-type phases were determined. Differential thermal analysis and high-temperature XRD (fig. S9) indicated that the solid-solution series is thermally stable up to 400°C.

### Ionic conductivity

The series of monophasic LSi$M_\delta$PSBrO samples showed conductivities higher than 10 mS cm$^{-1}$ at room temperature, even in the compressed powder state with grain boundary resistance (table S4 and fig. S10). The intracrystal conductivity was further evaluated for the composition at $M$ = Ge and $\delta$ = 0.4 in LSi$M_\delta$PSBrO (i.e., $Li_{9.54}[Si_{0.6}Ge_{0.4}]_{1.74}P_{1.44}S_{11.1}Br_{0.3}O_{0.6}$, abbreviated as LSiGePSBrO hereafter), because the LSiGePSBrO phase formed with the high phase purity and showed nearly maximum conductivity values. Figure 3 shows the Arrhenius plots in the temperature range of −50°C to 60°C for hot-pressed LSiGePSBrO (see figs. S11 to 13 for details of measurements). The bulk conductivity at room temperature was determined to be 32 mS cm$^{-1}$, which is approximately three times higher than that of the original LGPS and is the highest conductivity among $Li^+$ conductors reported to date. It should be noted that at −10°C, LSiGePSBrO still exhibits a high conductivity of 9 mS cm$^{-1}$, which is comparable to that of LGPS at room temperature. These extremely high conductivities can potentially facilitate the efficient use of active materials in all-solid-state cells particularly with thick cathode configurations.

### Application to thick battery electrode

To demonstrate the potential of highly conductive LSiGePSBrO in ASSLBs, it was used in a heavily loaded cathode by mixing with $LiNbO_3$-coated $LiCoO_2$ (LNO-LCO) as the active material. A solid-state cell was prepared by loading the cathode mixture at 245 mg cm$^{-2}$, which corresponds to an areal theoretical capacity of 23.5 mAh cm$^{-2}$ and a thickness of 800 μm. Figure 4A shows the discharge curves at a current density of 0.587 mA cm$^{-2}$ (0.025 C rate) in the temperature range of 25°C to −10°C. At 25°C, the cell showed an areal discharge capacity of 22.7 mAh cm$^{-2}$ with 97% utilization efficiency for the input LCO at the cutoff voltage of 4.25 V versus $Li^+$/Li. This observed areal discharge capacity is compared with others for ASSLBs (Fig. 4B) (*4*, *14*, *27*–*30*). More than 92% areal discharge capacity was reattained after the cycle test over 100 days, indicating no significant decomposition of LSiGePSBrO in the cathode at 4.25 V versus $Li^+$/Li (figs. S14 and S15). When the temperature decreased, the discharge capacity gradually decreased and the overpotential increased. The observed overpotential is attributed to the gradual decrease in the ionic conductivity of LSiGePSBrO at lower temperatures. Nevertheless, the discharge capacity at −10°C is still 75% of that at 25°C. Compared with recently reported heavily loaded all-solid-state cells (*4*, *27*), the present cell with a higher cathode weight delivered better capacity retention at low temperatures (Fig. 4C). Combined with lithium metal anode, the cathode composite with LSiGePSBrO discharged at a large current density of 14 mA cm$^{-2}$ at 60°C (fig. S16). The charge–discharge operation was confirmed even for an electrode 1-mm thick (fig. S17).

In conclusion, a series of solid electrolytes with the composition of $Li_{9.54}[Si_{1-\delta}M_\delta]_{1.74}P_{1.44}S_{11.1}Br_{0.3}O_{0.6}$ ($M$ = Ge, Sn; $0 \leq \delta \leqq 1$) was designed to achieve high configuration entropy while preserving their crystal structure with a superionic conduction pathway. The bulk conductivity of monophasic LSiGePSBrO ($M$ = Ge, $\delta$ = 0.4) was found to be 32 mS cm$^{-1}$ at 25°C. Theoretical calculations and structural analyses suggested that even the small degree of

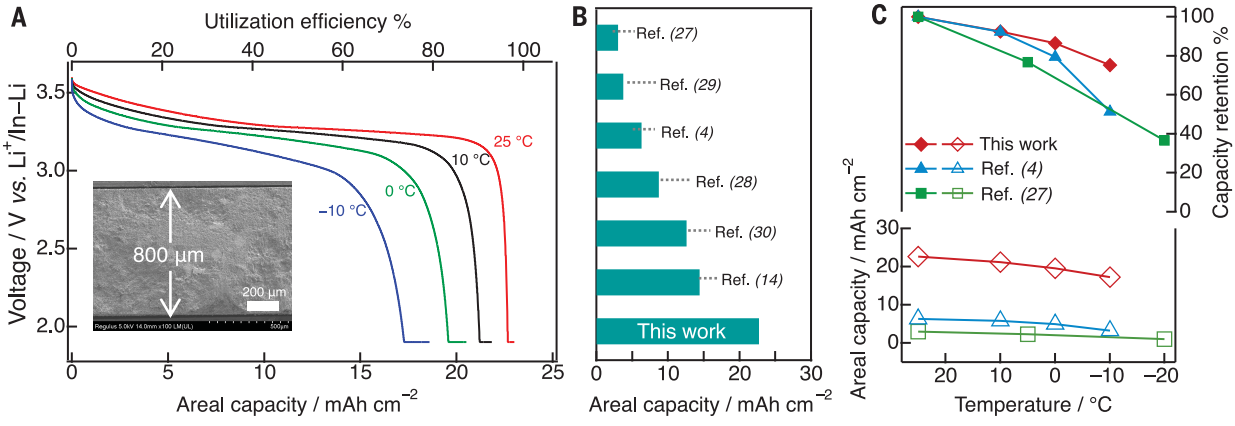



**Fig. 4. Performance of all-solid-state cells.** (**A**) Discharge curves for the thick electrode presented herein (~800 μm) under 0.587 mA cm$^{-2}$ (0.025 C) at different temperatures ranging from 25°C to −10°C. Inset: cross sectional scanning electron microscopy image of the cathode composite pellet with a capacity loading of 23.5 mAh cm$^{-2}$. (**B**) Comparison of areal capacities at 25°C for present and reported all-solid-state cells. The applied current density was between 0.2 and 0.68 mA cm$^{-2}$. (**C**) Comparison of areal capacities (lower) and discharge capacity retention rate (upper) at 25°C and lower temperatures.

chemical substitution in LSiGePSBrO could flatten the energy barrier for ion migration, thereby rationalizing the observed conductivity enhancement in this phase. An all-solid-state cell with a heavily loaded cathode (thickness 800 μm) incorporating LSiGePSBrO as the catholyte demonstrated a discharge capacity of 22.7 and 17.3 mAh $cm^{-2}$ at 25 and −10°C, corresponding to the utilization efficiency of 97 and 73% for the active material, respectively. This study highlights the significance of a high conductivity in the charge–discharge performance of ASSLBs using a thick cathode configuration. The design rule presented herein may accelerate the exploration of superionic conductors.

## REFERENCES AND NOTES


1. T. Famprikis, P. Canepa, J. A. Dawson, M. S. Islam, C. Masquelier, *Nat. Mater.* **18**, 1278–1291 (2019).
2. L. Ye, X. Li, *Nature* **593**, 218–222 (2021).
3. S. Randau *et al*., *Nat. Energy* **5**, 259–270 (2020).
4. Y. G. Lee *et al*., *Nat. Energy* **5**, 299–308 (2020).
5. Y. Kato *et al*., *Nat. Energy* **1**, 16030 (2016).
6. A. Hammami, N. Raymond, M. Armand, *Nature* **424**, 635–636 (2003).
7. Q. Wang, J. Sun, X. Yao, C. Chen, *Thermochim. Acta* **437**, 12–16 (2005).
8. N. Kamaya *et al*., *Nat. Mater.* **10**, 682–686 (2011).
9. M. A. Kraft *et al*., *J. Am. Chem. Soc.* **140**, 16330–16339 (2018).
10. Y. Seino, T. Ota, K. Takada, A. Hayashi, M. Tatsumisago, *Energy Environ. Sci.* **7**, 627–631 (2014).
11. H.-J. Deiseroth *et al*., *Angew. Chem. Int. Ed.* **47**, 755–758 (2008).
12. T. Yajima *et al*., *J. Mater. Chem.* **9**, 11278–11284 (2021).
13. Y. Wang *et al*., *Nat. Mater.* **14**, 1026–1031 (2015).
14. Y. Kato *et al*., *J. Phys. Chem. Lett.* **9**, 607–613 (2018).
15. A. Kubanska, L. Castro, L. Tortet, M. Dollé, R. Bouchet, *J. Electroceram.* **38**, 189–196 (2017).
16. S. Hull, D. A. Keen, D. S. Sivia, P. Berastegui, *J. Solid State Chem.* **165**, 363–371 (2002).
17. T. Takahashi, R. Kanno, Y. Takeda, O. Yamamoto, *Solid State Ion.* **3–4**, 283–287 (1981).
18. D. O. Raleigh, *J. Appl. Phys.* **41**, 1876–1877 (1970).
19. B. B. Owens, G. R. Argue, *Science* **157**, 308–310 (1967).
20. T. Minami, A. Hayashi, M. Tatsumisago, *Solid State Ion.* **177**, 2715–2720 (2006).
21. Y. Deng *et al*., *ACS Appl. Mater. Interfaces* **9**, 7050–7058 (2017).
22. D. Bérardan, S. Franger, A. K. Meena, N. Dragoe, *J. Mater. Chem. A* **4**, 9536–9541 (2016).
23. N. Dragoe, D. Bérardan, *Science* **366**, 573–574 (2019).
24. R. D. Shannon, *Acta Crystallogr. A* **32**, 751–767 (1976).
25. M. A. Kraft *et al*., *J. Am. Chem. Soc.* **139**, 10909–10918 (2017).
26. S. V. Patel *et al*., *Chem. Mater.* **33**, 1435–1443 (2021).
27. D. H. S. Tan *et al*., *Science* **373**, 1494–1499 (2021).
28. C. Doerrer *et al*., *ACS Appl. Mater. Interfaces* **13**, 37809–37815 (2021).
29. L. Zhou *et al*., *Nat. Energy* **7**, 83–93 (2022).
30. S. Wang *et al*., *Adv. Energy Mater.* **11**, 2101370 (2021).


## ACKNOWLEDGMENTS

The synchrotron radiation experiments were performed at the BL02B2 beamline of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposals 2020A1645 and 2021B1776). The neutron diffraction experiments were conducted as part of a project approved by the Neutron Scattering Program Advisory Committee of the Institute of Materials Structure Science (IMSS) and KEK (Proposal 2019S10). **Funding:** This research was partly supported by ALCA-SPRING [Grant JPMJAL1301], and the Program on Open Innovation Platform with Enterprises, Research Institute and Academia (OPERA) Program [JPMJOP1862] from the Japan Science and Technology Agency (JST). R.K. acknowledges support from the SOLiD-EV project [JPNP18003] of the New Energy and Industrial Technology Development Organization (NEDO). M.H. and K.S. acknowledge support from a Grant-in-Aid for Scientific Research on Innovative Area [19H05785]. **Author contributions:** K.S. and M.H. conceived the research idea. R.K. supervised the project. Y.L. and S.S. synthesized the samples. N.M., T.S., and T.K. performed the diffraction measurements. Hans. K., K.N., X.S., and Hanv. K. characterized samples and analyzed the data. T.M. performed ab initio simulations. The manuscript was drafted by Y.L. and S.H. and was written through contributions of all authors. All authors have given approval to the final version of the manuscript. **Competing interests:** Y.L., S.H., and R.K. hold patents related to this work (Japanese unexamined patent application publication 2022-030426). **Data and materials availability:** All data are available in the main text or the supplementary materials.
**License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add7138
Materials and Methods
Figs. S1 to S17
Tables S1 to S4
References (*31*–*69*)

Submitted 9 July 2022; accepted 25 May 2023
10.1126/science.add7138

PLANT SCIENCE

# Cortical polarity ensures its own asymmetric inheritance in the stomatal lineage to pattern the leaf surface

Andrew Muroyama[1,2]*, Yan Gong[1]†, Kensington S. Hartman[2], Dominique C. Bergmann[1,3]*

Asymmetric cell divisions specify differential cell fates across kingdoms. In metazoans, preferential inheritance of fate determinants into one daughter cell frequently depends on polarity-cytoskeleton interactions. Despite the prevalence of asymmetric divisions throughout plant development, evidence for analogous mechanisms that segregate fate determinants remains elusive. Here, we describe a mechanism in the *Arabidopsis* leaf epidermis that ensures unequal inheritance of a fate-enforcing polarity domain. By defining a cortical region depleted of stable microtubules, the polarity domain limits possible division orientations. Accordingly, uncoupling the polarity domain from microtubule organization during mitosis leads to aberrant division planes and accompanying cell identity defects. Our data highlight how a common biological module, coupling polarity to fate segregation through the cytoskeleton, can be reconfigured to accommodate unique features of plant development.

Asymmetric divisions establish differential cell identities by positioning daughter cells relative to fate-enforcing extrinsic signals or by segregating intrinsic fate determinants (*1*, *2*). The development of many plant organs depends on asymmetric divisions that place daughter cells in proximity to neighbor-derived signals such as mobile transcription factors (*3*, *4*), hormones (*5*–*7*), or small peptides (*8*, *9*). However, whether plants have mechanisms that ensure preferential inheritance of fate regulators remains unresolved.

To investigate how plant cells couple asymmetric divisions to identity specification, we monitored the cell division and differentiation dynamics of stomatal precursors in the *Arabidopsis* leaf epidermis (Fig. 1A). In this lineage, flexible asymmetric divisions in morphologically heterogeneous early lineage cells create and pattern stomata (cellular valves that mediate plant-atmosphere gas exchange) by priming divergent developmental trajectories in the two daughters. The smaller meristemoid will eventually give rise to the paired guard cells that comprise a stoma, and the larger stomatal lineage ground cell (SLGC) will expand to become a pavement cell.

All asymmetric cell divisions within the lineage are preceded by the formation of a plasma membrane–associated polarity complex defined by BREAKING OF ASYMMETRY IN THE STOMATAL LINEAGE (BASL) (*10*) and BREVIS RADIX family (BRXf) (*11*) proteins (Fig. 1, A and B). Before mitosis, the BASL/BRXf crescent recruits another polarly localized protein, POLAR (*12*), which in turn enriches GLYCOGEN SYNTHASE KINASE3 (GSK3)-like kinases at the cortex to promote asymmetric divisions (*13*), and then directs nuclear migration to bias the division site (*14*). After it is inherited by the SLGC through asymmetric division, cortical BASL/BRXf enhances MITOGEN ACTIVATED PROTEIN KINASE (MAPK) signaling to suppress meristemoid identity (*13*, *15*). Despite its central role in coordinating asymmetric division entry and daughter cell identity after division (*16*), how BASL/BRXf asymmetric inheritance is regulated is unknown (Fig. 1B).

## Stomatal lineage cells break common division orientation rules

We performed time-lapse imaging of developing cotyledons harboring markers for nuclei [R2D2 (*17*)], the plasma membrane (ML1p::mCherry-RCI2A), and the polarity crescent (BRXL2p::BRXL2-YFP). In agreement with previous observations, all asymmetric divisions resulted in singular inheritance of the BRXL2 crescent. Close analysis of these cells, however, revealed two asymmetric division subclasses that were defined by their division planes. The majority (73% of divisions) divided along the calculated shortest distance that intersected opposing cell walls at the site of the nucleus [i.e., a small angle between the calculated shortest wall and the division plane (Δθ°); see the materials and methods] (Fig. 1, C to F), thus following the expectations set by the observed division planes in many plant cell types (*18*, *19*). For this class of asymmetric divisions, polarity-directed nuclear migration (*14*) coupled with minimization of the division plane accurately predicted the final division site.

Division planes in the second class of asymmetric divisions (27% of divisions) deviated significantly from the calculated shortest wall (Fig. 1, D, E, and G), suggesting that additional inputs control the orientation of these e lineage divisions. Stomatal lineage divis are not unique in breaking the shortest wall rule, but other occurrences of them during *Arabidopsis* development are related to broad, extrinsic influences such as tissue mechanics or hormone signaling (*20*, *21*). Discrete, cell-autonomous mechanisms that tune division orientation have not been described. The morphological heterogeneity of stomatal precursors was well represented within both asymmetric division subclasses (Fig. 1, F and G), suggesting that unique geometric features do not define asymmetric division subtypes. Instead, asymmetrically dividing cells specifically bypassed the calculated shortest division plane (large Δθ°) when that wall was predicted to intersect the plasma membrane within the cortical polarized site (Fig. 1G). If only the nonpolarized membrane was considered permissive for division plane placement, then these asymmetric divisions continued to follow the shortest wall rule (fig. S1). This correlation suggested that polarized BASL/BRXf may be a cell-intrinsic cue capable of constraining potential orientations to control its own asymmetric inheritance.

Next, we tested whether cell polarity is necessary to stratify the two asymmetric division classes by tracking progenitor divisions in *basl* mutants (*basl* 35Sp::PIP2A-RFP ML1p::H2B-YFP). Loss of cellular polarity in *basl* collapsed the two asymmetric division classes into one (Fig. 1, H and I) that varied significantly from the total wild-type asymmetric divisions (Kolmogorov-Smirnov test, $P = 0.0005$). *basl* divisions did not differ significantly from wild-type asymmetric divisions without polarity conflict ($P = 0.1568$), indicating that *basl* divisions follow the shortest wall rule. Therefore, the polarized BASL/BRXf domain is required to override default division patterns during formative asymmetric divisions.

## BASL influences preprophase band position

To determine the basis of this control, we examined the cortical microtubule structures that play essential roles during division orientation in plant cells (*22*). TAN1p::CFP-TAN1 foci, which mark the cortical division zone (*23*), never appear within the BASL/BRXf domain, suggesting that BASL/BRXf operate at an early step during division orientation (fig. S2). We found that the preprophase band of microtubules, which is the first marker of the eventual cortical division site, never formed within the polarity domain (Fig. 1, J to M). An analysis of the Δθ° between the preprophase band and the calculated shortest wall showed a similar bifurcation of asymmetric divisions into two classes. The majority (66%) had preprophase bands that closely aligned to the predicted shortest wall, whereas preprophase bands in the second class (34%) deviated

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA. [2]Division of Biological Sciences, Department of Cell and Developmental Biology, University of California San Diego, La Jolla, CA 92093, USA. [3]Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA.
*Corresponding author. Email: bergmann@stanford.edu (D.C.B.); amuroyama@ucsd.edu (A.M.)
†Present address: Department of Organismic & Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.

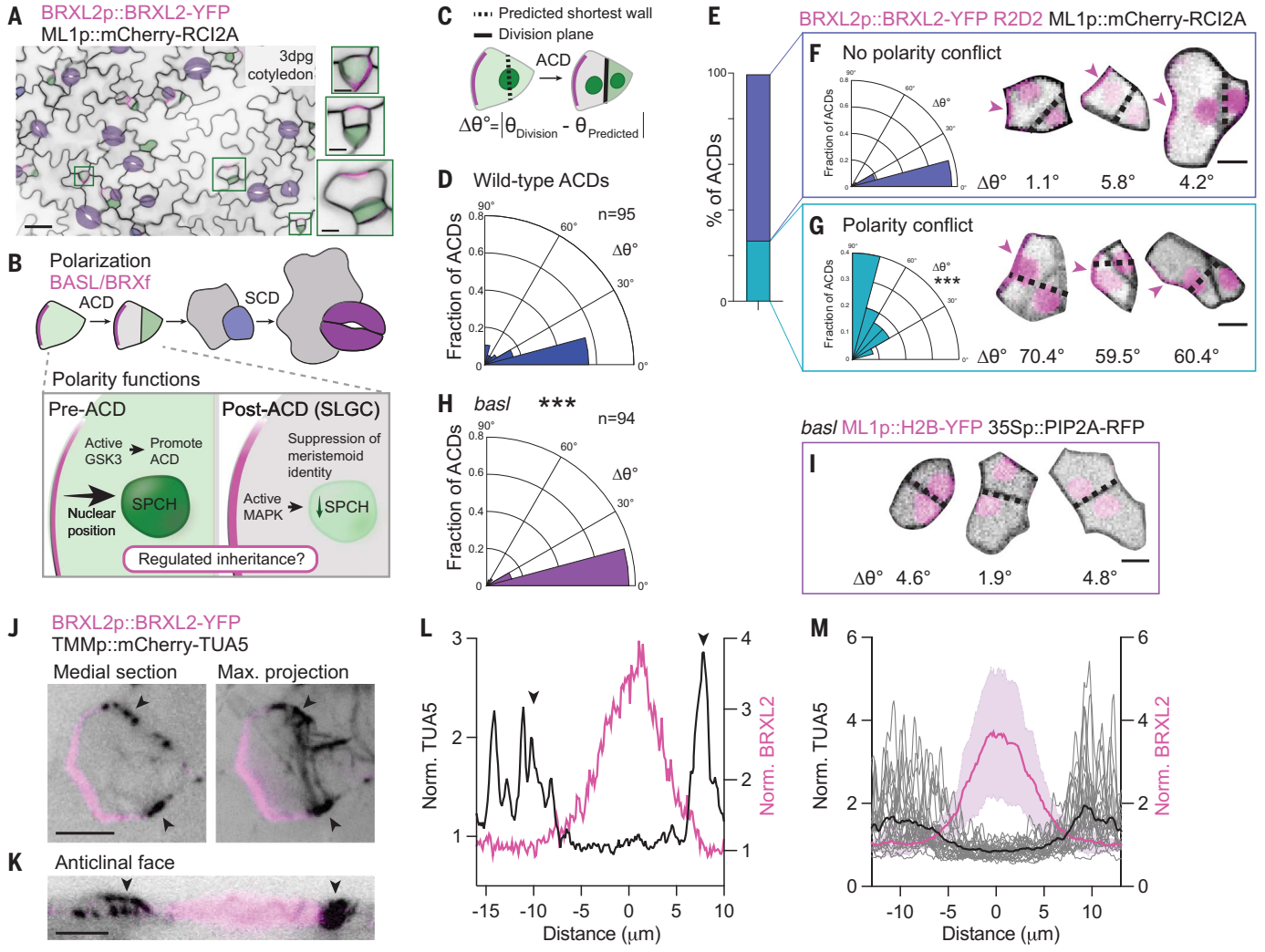



**Fig. 1. The *Arabidopsis* stomatal lineage polarity domain overrides default division rules during asymmetric cell division.** (**A**) Left: image of the abaxial epidermis from a 3 dpg BRXL2p::BRXL2-YFP (magenta) ML1p::mCherry-RCI2A (black) cotyledon. Cells have been pseudocolored according to the color scheme below. Scale bar, 25 μm. Right: selected polarized cells highlighting the morphological diversity of cells in this lineage. Scale bar, 5 μm. (**B**) Stomata (purple) are created through coupled divisions (asymmetric, ACD, and symmetric, SCD) and fate transitions. Top: illustration of the major fate transitions within the *Arabidopsis* stomatal lineage. Meristemoids are shown in green and the guard mother cell in blue. Bottom: illustration of the known functions for the BASL/BRXf domain. Before division in SPEECHLESS-positive ($SPCH^+$) cells, cortical BASL/BRXf promotes ACD and instructs nuclear migration. After division in SLGCs, inherited BASL/BRXf/YODA initiates a MAPK cascade, leading to suppression of SPCH. Mechanisms controlling singular polarity domain inheritance are needed. (**C**) The assay used to quantify Δθ°. This assay was used for the data shown in (D) to (I). For additional details, see the materials and methods. (**D**) Quantification of Δθ° during wild-type ACDs. *n* = 95 cells. (**E**) Percentage of total wild-type ACDs in which polarity did (27%) or did not (73%) conflict with the predicted shortest wall. (**F** and **G**) Left: distribution of Δθ° in ACDs in which the predicted shortest wall did (G) or did not (F) conflict with the polarity domain. Right: three examples of ACDs from their respective classes with associated Δθ°s. The black dotted line marks the calculated shortest wall, and the magenta arrow indicates the polarity domain. Scale bar, 5 μm. $P < 0.0001$, Kolmogorov-Smirnov test comparing the Δθ°s distributions in the two ACD classes. (**H**) Quantification of Δθ° during presumed meristemoid divisions in *basl*. $P = 0.0005$, Kolmogorov-Smirnov test with WT ACDs. *n* = 94 cells. (**I**) Examples of three early-lineage divisions in *basl* with associated Δθ°s. The black dotted line marks the predicted shortest wall. Scale bar, 5 μm. (**J**) Representative images of a medial (left) and maximum projection (right) view of preprophase band (PPB) placement relative to the polarity domain in BRXL2p::BRXL2-YFP TMMp::mCherry-TUA5. Black arrows indicate the PPB. Scale bar, 5 μm. (**K**) Reslice of the cell in (J) showing microtubule and BRXL2 distribution along the anticlinal wall. Black arrows indicate the PPB. Scale bar, 5 μm. (**L**) Line scan along the cell cortex of the asymmetrically dividing cell shown in (J). The arrows indicate the position of the PPB. (**M**) Microtubule distribution (gray lines) in PPB-forming cells (*n* = 26 cells, average is shown in black), aligned to the midpoints of the BRXL2 crescents (magenta line shows the mean ± SD).

significantly from the shortest distance (fig. S3). In this second class, preprophase bands did not align with the predicted shortest wall when they bisected the polarity domain, and polarity was required for preprophase band realignment away from the shortest wall (fig. S3). Together, these data indicated that the BASL/BRXf polarity crescent might orient divisions by controlling preprophase band placement.

To test this hypothesis, we generated lines to monitor BRXL2 inheritance in the *trm678* mutant (*trm678* BRXL2p::BRXL2-YFP ML1p::mCherry-RCI2A), which does not form preprophase bands (*24*). In contrast to wild-type

asymmetric divisions, in which BRXL2 was inherited by a single daughter cell, new cell walls frequently bisected the polarity site in *trm678* (32% of divisions) (Fig. 2, A and B), showing that the preprophase band is required to ensure complete inheritance of polarized BASL/BRXf. Next, we monitored cell fate outcomes following incorrect BASL/BRXf inheritance by tracking progression through the stomatal lineage using a fluorescent reporter of MUTE (*25*), a transcription factor that establishes guard mother cell identity. *trm678* asymmetric divisions in which BRXL2 was correctly inherited by a single daughter cell showed normal lineage progression. MUTE expression was detectable in the smaller cell after the division, and all tracked MUTE$^+$ cells in *trm678* became paired guard cells (fig. S4). By contrast, *trm678* cells in which cortical BRXL2 was bisected by the nascent division plane tended to generate daughters that both inherited cortical BRXL2, never transitioned to MUTE$^+$ cells, and failed to become pavement cells (Fig. 2C). In agreement with these tracking data, 7 d postgermination (dpg) *trm678* cotyledons had fewer stomata and a mispatterned epidermis (Fig. 2, D to F). We did not observe any alterations in the relationship between microtubules and BRXL2 in interphase *trm678* cells (fig. S4, D to F). Therefore, we conclude that the preprophase band serves as an essential link between the polarity domain and division orientation to regulate stomatal identity.

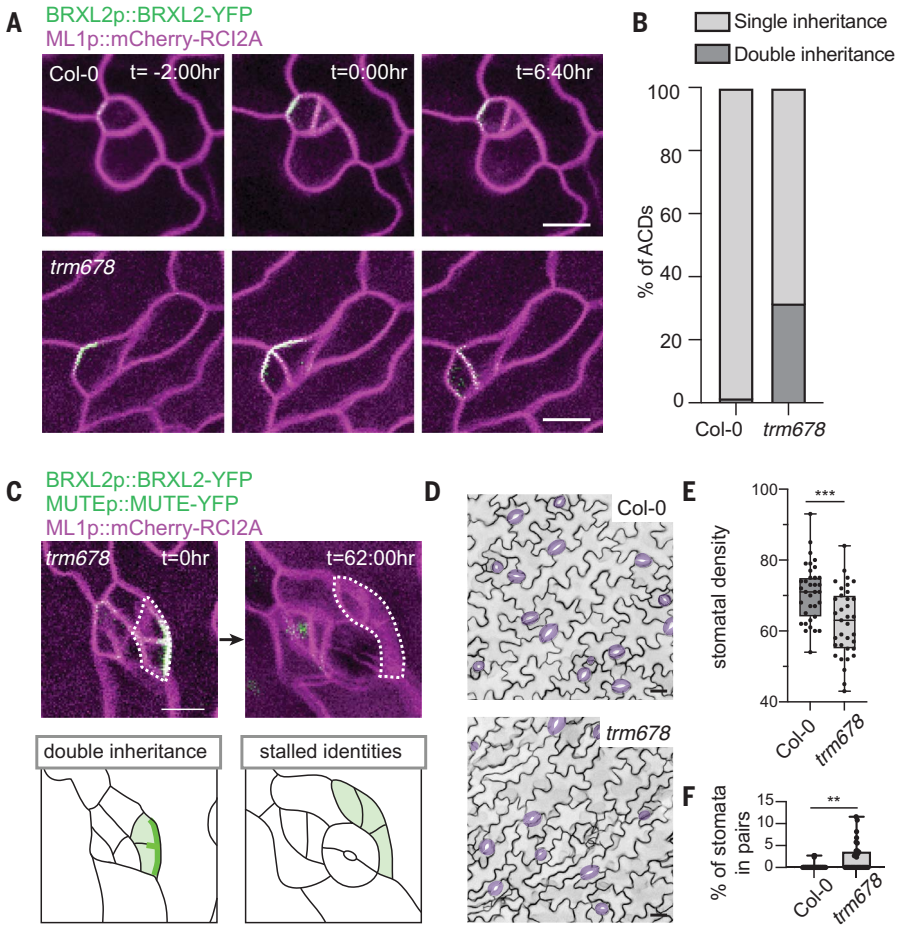



**Fig. 2. Singular inheritance of the BASL/BRXf polarity domain through preprophase band placement is required for stomatal patterning.** (**A**) Stills from representative time-lapse movies showing cell division orientation in Col-0 (top) and *trm678* (bottom) relative to the BRXL2-YFP-marked cortical polarity domain. Scale bars, 10 μm. (**B**) Quantification of Col-0 ($n$ = 112) and *trm678* ($n$ = 116) ACDs with singular inheritance (BRXL2 inherited by only one daughter) or double inheritance (BRXL2 inherited by both daughter cells). (**C**) Representative time-course images with associated cartoons showing an instance of double inheritance and cell fate stalling 62 hours later from a *trm678* seedling. Scale bar, 25 μm. (**D**) Representative images of the epidermis of 7 dpg Col-0 and *trm678* cotyledons expressing ML1p::mCherry-RCI2A. Stomata are pseudocolored purple. Scale bars, 25 μm. (**E**) Quantification of stomatal densities (number of stomata per 581.82 μm × 581.82 μm area) in 7 dpg Col-0 and *trm678* cotyledons (35 seedlings each). $P$ = 0.0002, unpaired $t$ test. (**F**) Quantification of the percentage of paired stomata per 581.82 μm × 581.82 μm area in 7 dpg Col-0 and *trm678* cotyledons (35 seedlings each). $P$ = 0.0012, unpaired $t$ test.

### Cortical BASL domains are locally depleted of stable microtubules

How does the BASL/BRXf crescent influence preprophase band establishment? By creating a stomatal lineage–specific microtubule reporter line (TMMp::mCherry-TUA5), we could analyze microtubules and the BRXL2 polarity domain along anticlinal walls with high resolution (Fig. 3A). Unexpectedly, anticlinal microtubules were strongly depleted from the plasma membrane within the polarized domain, even in interphase SLGCs (Fig. 3, A to C). We confirmed that the same microtubule depletion zone occurred when using a second polarity reporter, BASL, and in stomatal lineage cells of true leaves (fig. S5). POLAR, which shows overlapping but distinct localization from BASL/BRXf (*12*), co-localizes with microtubules outside of the BASL/BRXL2 domain (fig. S5), indicating that microtubule depletion is correlated specifically with BASL/BRXf and is not a generalized activity of polarized proteins in the stomatal lineage.

BASL/BRXf could either locally deplete microtubules or opportunistically polarize to already microtubule-poor regions. To distinguish between these possibilities, we performed two analyses. First, we examined cortical microtubule distribution before BASL polarization and found no microtubule-depleted region (fig. S6). Second, we compared microtubule distribution in wild-type and polarity-defective SLGCs and found that microtubule distribution was more homogeneous in *basl* and *brx-quad (11)*, which abrogate polarity, and in a line in which addition of a myristoylation signal [BASLp::MYR-BRX-YFP (*11*)] renders BRX localization largely uniform at the cortex (Fig. 3, D to G, and fig. S7). We also followed unmanipulated SLGCs as they lost polarized BASL/BRXf several hours after cell division. Our time course analysis showed that anticlinal microtubules reappeared within previously polarized regions in mature SLGCs (fig. S7E). From these data, we conclude that BASL/BRXf polarity creates and is required to maintain local microtubule loss at the plasma membrane.

Mutual inhibition by opposing plasma membrane–associated domains can drive polarization, as in the conserved PAR networks in animals (*26*) or in a recently described polarity system in the monocot *Brachypodium distachyon* (*27*). Because our data raised the possibility that microtubules and cortical BASL/BRXf could operate in an analogous manner and inhibit the spread of each other, we tested whether altering microtubule distribution affected the stomatal lineage polarity domain.

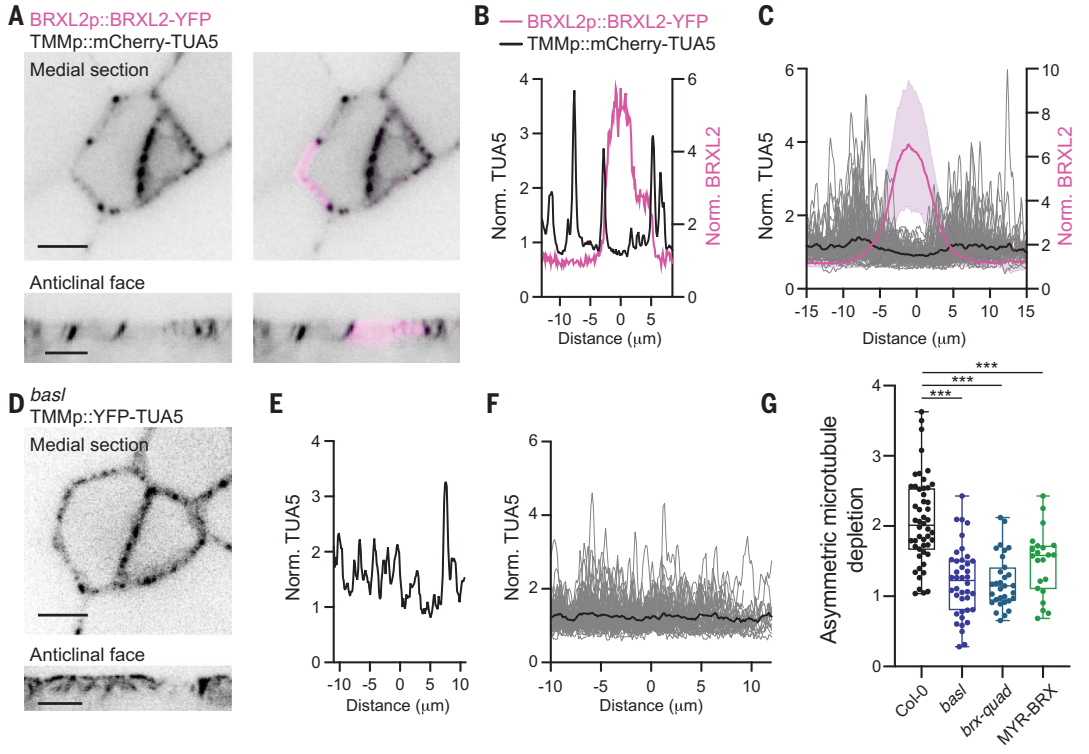



**Fig. 3. The BASL/BRXL2 domain locally depletes cortical microtubules from the plasma membrane.** (**A**) Top: representative medial section of a polarized SLGC and associated meristemoid in BRXL2p::BRXL2-YFP TMMp::mCherry-TUA5. Bottom: reslice showing microtubule and BRXL2 distribution along the anticlinal face of the same cell. Scale bars, 5 μm. (**B**) Line scan along the cell periphery of the SLGC shown in (A) showing the normalized BRXL2 and TUA5 fluorescence intensities. (**C**) Microtubule distribution in SLGCs (gray lines) ($n$ = 72 cells) aligned to the midpoints of the BRXL2 crescents. Magenta line corresponds to the mean signal ± SD. Black line is the average of the plotted microtubule signals. (**D**) Top: representative medial section of stomatal lineage cells in *basl* TMMp::YFP-TUA5. Bottom: reslice showing microtubule distribution along the anticlinal face of the same cell. Scale bars, 5 μm. (**E**) Line scan along the cell cortex of the larger daughter cell in (D) showing the normalized TUA5 fluorescence intensity. (**F**) Microtubule distribution in *basl* SLGCs (gray lines) ($n$ = 58 cells). (**G**) Asymmetric microtubule depletion scores calculated from integrated fluorescence intensities in polarized and unpolarized cortical domains (see the materials and methods) in wild-type ($n$ = 50), *basl* ($n$ = 40), *brx-quad* ($n$ = 31), and BASLp::MYR-BRX ($n$ = 22) SLGCs. ***$P$ < 0.0001, one-way ANOVA with Tukey's post hoc test.

In agreement with previous results (*28*, *29*), we found that microtubules were not necessary for the formation of a polarized BASL/BRXf domain (fig. S8A). However, quantification revealed a slight but significant spread of the polarity domain along the anticlinal wall in the absence of microtubules (fig. S8B). Short plasmolysis treatments, which substantially disrupt cortical microtubule distribution, similarly altered polarity boundaries and polarity domain size without complete depolarization (fig. S8C). Therefore, microtubules shape BASL/BRXf domain boundaries but are dispensable for polarity itself. This interaction shares marked similarities with microtubule-mediated sculpting of ROP GTPase domains in nondividing cells of the xylem and in trichomes (*30*–*32*).

### Microtubule dynamics are locally altered within polar domains

How do cortical BASL/BRXf locally deplete cortical microtubules? Because of the technical challenges associated with monitoring dynamic microtubule behavior along the anticlinal walls of meristemoids, we created a heterologous system in which we could track microtubule dynamics coincident with the BASL polarity domain. By introgressing a ubiquitous microtubule reporter, 35Sp::mCherry-TUA5, into a line expressing a hyperactive version of BASL capable of rescuing the *basl* phenotype, 35Sp::GFP-BASL-IC [hereafter referred to as BASL$^{ectopic}$ (*10*)], we could monitor microtubule organization relative to BASL in the hypocotyl epidermis. BASL$^{ectopic}$ locally depleted microtubules along anticlinal walls in the hypocotyl epidermis, as in the stomatal lineage (fig. S9), demonstrating that this ectopic system recapitulates the molecular interactions found in the leaf epidermis.

BASL$^{ectopic}$ domains extended onto the apical surfaces of hypocotyl epidermal cells and locally depleted cortical microtubules (Fig. 4, A and C), allowing us to observe the BASL-mediated effects on microtubules with a precision not possible within the stomatal lineage. Increased microtubule severing has been identified as a key reorganizer of cortical microtubule arrays during several developmental transitions (*33*, *34*). However, because severing preferentially occurs at microtubule crossover sites (*35*, *36*) and there were few crossovers in microtubule-depleted BASL$^{ectopic}$ regions, severing was largely suppressed within BASL$^{ectopic}$ domains (fig. S10, A and B). Tracking of microtubule minus ends within BASL$^{ectopic}$ also indicated that local microtubule depletion was not caused by decreased minus-end stability (fig. S10C). Instead, we found that BASL$^{ectopic}$ had two effects on microtubule plus ends. First, plus-end polymerization and depolymerization rates were significantly suppressed within BASL$^{ectopic}$ (Fig. 4, E and F, and fig. S10D). Second, we observed that microtubule plus ends rapidly underwent catastrophe upon entering the BASL$^{ectopic}$ domain (Fig. 4, B and D). Increased catastrophe rates often led to complete loss of the microtubule, reestablishing the microtubule depletion zone.

We used two independent approaches to validate that our BASL$^{ectopic}$ findings in the hypocotyl reflect BASL-microtubule interactions within the stomatal lineage. First, we performed time-lapse imaging of stomatal progenitors in BRXL2p::BRXL2-YFP TMMp::mCherry-TUA5 seedlings and observed transient microtubules within the polarity domain that were rapidly depolymerized (fig. S11A). Second, to monitor growing plus ends with higher precision along the anticlinal wall, we generated a stomatal lineage–specific END BINDING PROTEIN 1b (EB1b) reporter, TMMp::EB1b-mCherry, and introgressed it into the BRXL2 reporter line. Fewer EB1b puncta were observed within the native polarity domain in SLGCs than in nonpolarized regions of the same cells (fig. S11, B to D). Therefore, our analyses of microtubule dynamics in native and heterologous systems

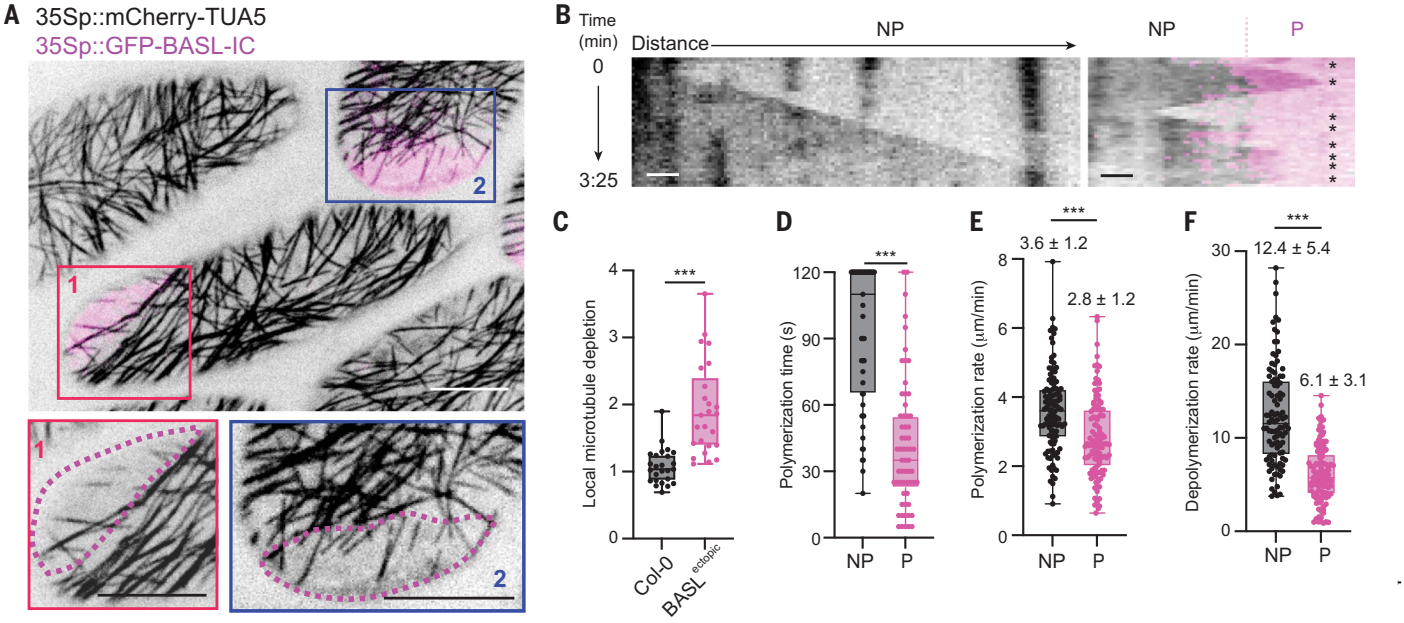



**Fig. 4. Polarized BASL destabilizes microtubule plus ends to locally bias array organization.** (**A**) Representative images of the apical surfaces in epidermal cells of 35Sp::GFP-BASL-IC 35Sp::mCherry-TUA5 hypocotyls. The boxed regions below highlight the microtubule organization within the polarized domains of two cells. Scale bars, 10 μm. (**B**) Kymographs showing microtubule plus-end dynamics within a nonpolarized region (left) and an ectopic polarity domain (right). Asterisks indicate microtubule catastrophes and rescues. Scale bar, 1 μm. (**C**) Local microtubule depletion within apical BASL$^{ectopic}$ domains and comparably sized random regions in control hypocotyls. The local microtubule depletion was derived using 35Sp::mCherry-TUA5 fluorescence (see the materials and methods). $n$ = 25 hypocotyl cells for each. $P < 0.0001$, unpaired $t$ test. (**D** to **F**) Quantification of microtubule plus-end dynamics within nonpolarized (NP) and polarized (P) apical domains. Growth persistence (D), polymerization rate (E), and depolymerization rate (F) were quantified. The numbers above the box-and-whisker plots are mean values ± SD. For all comparisons, unpaired $t$ tests were used ($P < 0.0001$).

reveal that the BASL/BRXf domain destabilizes microtubule plus ends to locally deplete them from the polarized region.

## Discussion

As encoders of spatial information, polarity domains are central regulators of asymmetric cell division in diverse plant tissues (*37*–*40*). Here, we provide evidence that the BASL/BRXf polarity domain robustly serves dual functions to orient asymmetric divisions and specify cell identity by controlling its own inheritance through negative interactions with the microtubule cytoskeleton. Polarity-microtubule interactions now emerge as a common theme to guide asymmetric inheritance of fate regulators during both metazoan and plant asymmetric divisions, albeit through fundamentally different mechanisms (fig. S12). In the canonical asymmetric cell division pathway in animal cells, the cortical polarity domain is responsible for localizing fate determinants to one pole and subsequently directing the division angle to ensure their singular and asymmetric inheritance by exerting pulling forces on astral microtubules (*41*, *42*). The model we advance here differs in several significant ways. First, the proposed mechanism uses core plant-specific mitotic structures without the need to invoke a role for astral microtubules, which are absent in plant spindles. Second, rather than ensuring its singular inheritance by precisely specifying the ultimate division plane, BASL/BRXf renders a region of the membrane unavailable as the division site. Third, this mechanism is uniquely suited to ensure asymmetric inheritance of key fate regulators within a morphologically heterogeneous population.

How does BASL-mediated polarity modulate microtubule dynamics? BASL and BRXf proteins both contain large, disordered regions (*43*), leading us to favor two general models. In the first, BASL and BRXf scaffold effectors that directly affect plus-end kinetics and potentially bind along the microtubule lattice to affect depolymerization rates. From our analyses in this work, we anticipate that such microtubule-associated effectors would be expressed throughout the cell cycle and in multiple tissues, which has complicated our ongoing efforts to identify them. In the second model, which does not invoke additional downstream factors, polarization through phase separation could tune the local physical properties of the membrane-adjacent cytoplasm. Such a mechanism can modulate microtubule dynamics in yeast (*44*) and might be hinted at by the dampening of microtubule dynamics at the polarity zone (fig. S10D).

Of the documented, polarity-mediated asymmetric divisions in plants, those that violate the shortest wall rule, such as those in the early *Arabidopsis* embryo or subsidiary mother cell divisions in *Zea mays*, may be the closest corollaries to the system presented here. Although BASL is both eudicot (*43*) and stomatal lineage specific, BRX family proteins participate in additional cellular decisions in *Arabidopsis* and are much more deeply conserved in the green lineage (*45*), hinting that other tissue-specific regulators could provide context specificity to a common polarity core. Further analysis of polarity will help to clarify whether this mechanism is shared across plant tissues and species or if it has evolved for the challenges associated with flexible patterning in the eudicot stomatal lineage.

## REFERENCES AND NOTES

1. B. Sunchu, C. Cabernard, *Development* **147**, dev167650 (2020).
2. Z. G. Venkei, Y. M. Yamashita, *J. Cell Biol.* **217**, 3785–3795 (2018).
3. K. Nakajima, G. Sena, T. Nawy, P. N. Benfey, *Nature* **413**, 307–311 (2001).
4. M. T. Raissig *et al.*, *Science* **355**, 1215–1218 (2017).
5. I. Casimiro *et al.*, *Plant Cell* **13**, 843–852 (2001).
6. A. Vatén, C. L. Soyars, P. T. Tarr, Z. L. Nimchuk, D. C. Bergmann, *Dev. Cell* **47**, 53–66.e5 (2018).
7. A. Cruz-Ramírez *et al.*, *Cell* **150**, 1002–1015 (2012).
8. K. Hara, R. Kajita, K. U. Torii, D. C. Bergmann, T. Kakimoto, *Genes Dev.* **21**, 1720–1725 (2007).
9. L. Hunt, J. E. Gray, *Curr. Biol.* **19**, 864–869 (2009).
10. J. Dong, C. A. MacAlister, D. C. Bergmann, *Cell* **137**, 1320–1330 (2009).

11. M. H. Rowe, J. Dong, A. K. Weimer, D. C. Bergmann, bioRxiv, 614636 (2019).
12. L. J. Pillitteri, K. M. Peterson, R. J. Horst, K. U. Torii, *Plant Cell* **23**, 3260–3275 (2011).
13. A. Houbaert *et al*., *Nature* **563**, 574–578 (2018).
14. A. Muroyama, Y. Gong, D. C. Bergmann, *Curr. Biol.* **30**, 4467–4475.e4 (2020).
15. Y. Zhang, P. Wang, W. Shao, J. K. Zhu, J. Dong, *Dev. Cell* **33**, 136–149 (2015).
16. Y. Gong *et al*., *Development* **148**, dev199919 (2021).
17. C. Y. Liao *et al*., *Nat. Methods* **12**, 207–210, 2, 210 (2015).
18. S. Besson, J. Dumais, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6294–6299 (2011).
19. B. E. Shapiro, C. Tobin, E. Mjolsness, E. M. Meyerowitz, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4815–4820 (2015).
20. P. Vaddepalli *et al*., *Curr. Biol.* **31**, 4946–4955.e4 (2021).
21. M. Louveaux, J. D. Julien, V. Mirabet, A. Boudaoud, O. Hamant, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E4294–E4303 (2016).
22. C. G. Rasmussen, A. J. Wright, S. Müller, *Plant J.* **75**, 258–269 (2013).
23. K. L. Walker, S. Müller, D. Moss, D. W. Ehrhardt, L. G. Smith, *Curr. Biol.* **17**, 1827–1836 (2007).
24. E. Schaefer *et al*., *Science* **356**, 186–189 (2017).
25. L. J. Pillitteri, D. B. Sloan, N. L. Bogenschutz, K. U. Torii, *Nature* **445**, 501–505 (2007).
26. B. Etemad-Moghadam, S. Guo, K. J. Kemphues, *Cell* **83**, 743–752 (1995).
27. D. Zhang *et al*., *eLife* **11**, e79913 (2022).
28. J. Chan, C. Mansfield, F. Clouet, D. Dorussen, E. Coen, *Curr. Biol.* **30**, 4999–5006.e3 (2020).
29. M. Bringmann, D. C. Bergmann, *Curr. Biol.* **27**, 877–883 (2017).
30. Y. Oda, H. Fukuda, *Science* **337**, 1333–1336 (2012).
31. Y. Sugiyama, M. Wakazaki, K. Toyooka, H. Fukuda, Y. Oda, *Curr. Biol.* **27**, 2522–2528.e4 (2017).
32. M. Yanagisawa, J. M. Alonso, D. B. Szymanski, *Curr. Biol.* **28**, 2459–2466.e4 (2018).
33. J. J. Lindeboom *et al*., *Science* **342**, 1245533 (2013).
34. R. Schneider *et al*., *Nat. Commun.* **12**, 669 (2021).
35. R. Wightman, S. R. Turner, *Plant J.* **52**, 742–751 (2007).
36. Q. Zhang, E. Fishel, T. Bertroche, R. Dixit, *Curr. Biol.* **23**, 2191–2195 (2013).
37. S. Yoshida *et al*., *Nat. Plants* **5**, 160–166 (2019).
38. R. Campos, J. Goff, C. Rodriguez-Furlan, J. M. Van Norman, *Dev. Cell* **52**, 183–195.e4 (2020).
39. H. N. Cartwright, J. A. Humphries, L. G. Smith, *Science* **323**, 649–651 (2009).
40. S. Yoshida *et al*., *Dev. Cell* **29**, 75–87 (2014).
41. T. Lechler, M. Mapelli, *Nat. Rev. Mol. Cell Biol.* **22**, 691–708 (2021).
42. T. Kiyomitsu, *Curr. Opin. Cell Biol.* **60**, 1–8 (2019).
43. I. Nir *et al*., *Curr. Biol.* **32**, 329–337.e5 (2022).
44. A. T. Molines *et al*., *Dev. Cell* **57**, 466–479.e6 (2022).
45. J. J. Ramalho, V. A. S. Jones, S. Mutte, D. Weijers, *Plant Cell* **34**, 174–192 (2022).


## ACKNOWLEDGMENTS

We thank D. Bouchez for sharing the *trm678* seeds; A. Lanctot for generating POLARp::POLAR-YFP; C. Rasmussen for sharing the TAN1p::CFP-TAN1 reporter seeds; N. Sharma for technical help generating the *trm678* reporter lines; E. Griffis at the Nikon Imaging Center at the University of California, San Diego, for assistance with several imaging experiments; and M. Estelle, A. Dickinson, and members of the Bergmann laboratory for feedback on the manuscript. **Funding:** A.M. was supported by a postdoctoral fellowship from the National Institutes of Health (grant F32 GM133102-01) and is currently supported by start-up funds from the University of California, San Diego. Y.G. was supported by funds from Stanford University and the Howard Hughes Medical Institute. K.S.H. is supported by funds from the University of California, San Diego to A.M. D.C.B. is an investigator of the Howard Hughes Medical Institute. **Author contributions:** A.M. and D.C.B. conceived of the study and designed the experiments. Y.G. acquired the time-lapse data for cell tracking in *trm678*. K.S.H. acquired and analyzed the *trm678* mutant phenotype data. A.M. performed all other experiments and analyses. A.M. and D.C.B. wrote the manuscript with feedback from Y.G. and K.S.H. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All newly generated materials in the manuscript are available upon reasonable request. All data are available in the manuscript and associated supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add6162
Materials and Methods
Figs. S1 to S12
References (*46*–*51*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

Submitted 22 June 2022; resubmitted 10 March 2023
Accepted 11 May 2023
10.1126/science.add6162

SOLAR CELLS

# Interface passivation for 31.25%-efficient perovskite/silicon tandem solar cells

Xin Yu Chin[1,2]*, Deniz Turkay[1], Julian A. Steele[3,4,5], Saba Tabean[6,7], Santhana Eswara[6], Mounir Mensi[8], Peter Fiala[1], Christian M. Wolff[1], Adriana Paracchino[2], Kerem Artuk[1], Daniel Jacobs[1], Quentin Guesnay[1], Florent Sahli[2], Gaëlle Andreatta[2], Mathieu Boccard[1], Quentin Jeangros[1,2], Christophe Ballif[1,2]

Silicon solar cells are approaching their theoretical efficiency limit of 29%. This limitation can be exceeded with advanced device architectures, where two or more solar cells are stacked to improve the harvesting of solar energy. In this work, we devise a tandem device with a perovskite layer conformally coated on a silicon bottom cell featuring micrometric pyramids—the industry standard—to improve its photocurrent. Using an additive in the processing sequence, we regulate the perovskite crystallization process and alleviate recombination losses occurring at the perovskite top surface interfacing the electron-selective contact [buckminsterfullerene ($C_{60}$)]. We demonstrate a device with an active area of 1.17 square centimeters, reaching a certified power conversion efficiency of 31.25%.

The deployment of photovoltaics (PV) can be accelerated by maximizing the electrical power produced per unit area because the cost distribution of a PV system is now dominated by the balance of system components (such as the mounting system, wiring, manpower, and inverters) rather than by the cost of the PV panels. This balance of system costs scales roughly with the installed area and favors PV technologies with a high power-to–panel area ratio. However, the record power conversion efficiency (PCE) of crystalline silicon (c-Si) solar cells of 26.8% (*1*) is approaching the theoretical limit of 29.5% (*2*). The only experimentally validated approach to overcome this PCE limitation under 1-sun illumination conditions consists of combining several complementary photoactive materials (meaning multiple junctions) in a single device (*3*). Among the different types of multijunction designs reported to date, the combination of c-Si with a metal halide perovskite in a tandem solar cell has been the focus of intense research efforts given the potential for both high PCE and low manufacturing costs (*4*).

Metal halide perovskites combine several key properties for effective multijunction PV, including a high absorption coefficient with a sharp absorption edge (*5*), ambipolar charge transport with long diffusion lengths (*6*, *7*), and a compositionally tunable bandgap energy ($E_g$) (*8*). A thin-film perovskite solar cell can be deposited directly on the front side of a c-Si cell to lower thermalization losses and extend the range of achievable PCE to >30% (*3*). The performance potential of monolithic two-terminal tandem architectures is illustrated by reported PCEs as high as 33.7% for 1-$cm^2$ illuminated areas (*1*). Most of the high-efficiency tandem cells reported to date use a Si wafer in which the front surface is mechanically or chemically polished (*9*) or features an adapted submicrometric texture smaller than the perovskite layer thickness (typically from 500 nm to 1 μm) (*10*–*13*). This planar or nanotextured front-side topography—wafers used in the PV industry are usually etched to form pyramids of a few micrometers in height—enables the deposition of pinhole-free perovskite films by standard solution-based processes. However, this modification comes at the expense of optical performance because the front side of the tandem is flat, and also when using a submicrometric Si texture, because the latter is planarized by the nonconformal solution-processed perovskite film. Consequently, these cell designs exhibit more reflection losses at the front of the tandem because of an absence of double-bounce effects (*14*). Overall, a pyramidal texture at the front of the tandem device limits reflection losses because it enables the absorption of reflected light in neighboring pyramids, whereas the presence of texture on both sides of the Si wafer improves the trapping of infrared light (*15*).

We previously reported a hybrid two-step deposition method combining thermal evaporation and spin coating to conformally coat the perovskite layer on micrometric Si p[yra]mids, resulting in perovskite/c-Si tandem[s in] which both rear and front sides were textured (*16*). Although these tandem cells had a high photocurrent thanks to the front-side pyramidal texture, nonradiative recombination losses were substantial. One challenge is that most top-surface passivation approaches reported to date are not directly applicable to micrometric textures because they involve depositing nanometric organic layers from liquid solutions (*17*–*19*). These processing routes usually yield nonconformal (incomplete) coatings on such surface textures. Building on our previous work (*16*), we demonstrate enhanced tandem performance using phosphonic acids in two different roles to passivate interfacial defects: first, as a hole-transport layer (HTL), [4-(3,6-dimethyl-9H-carbazol-9-yl)butyl]phosphonic acid (Me-4PACz) (*9*) and second, as a perovskite additive in the form of 2,3,4,5,6-pentafluorobenzylphosphonic acid (FBPAc) (*20*). Together with a micrometric texture standard in the c-Si PV industry and an optimization of the layers located at the front of the tandem for further optical gains, these passivation strategies mitigate voltage losses and result in tandem device efficiencies >30%, with an accredited value of 31.25%.

## Identification and mitigation of voltage losses in perovskite top cells

To assess loss pathways triggered by nonradiative recombination, we measured the photoluminescence quantum yield of different layer stacks representative of our p-i-n perovskite solar cells used in tandem cells. We quantified the internal voltage potential (in relative terms), or quasi–Fermi-level splitting (QFLS), created within the cesium formamidinium lead iodide bromide [$Cs_{0.18}FA_{0.82}Pb(I,Br)_3$] absorber under light excitation (Fig. 1A). Flat glass/indium tin oxide (ITO) substrates were first coated with a HTL, either 2,2′,7,7′-tetra(N,N-di-tolyl)amino-9,9-spiro-bifluorene (spiro-TTB), N4,N4,N4″,N4″-tetra([1,1′-biphenyl]-4-yl)-[1,1′:4′,1″-terphenyl]-4,400-diamine (TaTM), or Me-4PACz (fig. S1). These three materials were selected on the basis of their compatibility with textured surfaces. The first two can be thermally evaporated, and the latter can self-assemble on ITO to yield a conformal coating on textured surfaces (*21*). We then deposited a perovskite thin film using a hybrid two-step deposition method (*16*). Specifically, a CsBr/$PbI_2$ template was coevaporated on either flat glass substrates or glass/ITO/HTL stacks before spin coating an organo-halide solution consisting of FABr and FAI and, in some cases, with FBPAc. A thermal annealing step in an ambient atmosphere at 150°C converted the films into a perovskite phase (see the experimental section for details). In view of making tandems, a perovskite composition with an

[1]Photovoltaics and Thin-Film Electronics Laboratory (PV-Lab), Institute of Microengineering (IMT), Ecole Polytechnique Fédérale de Lausanne (EPFL), Neuchâtel, Switzerland. [2]Centre Suisse d'Électronique et de Microtechnique (CSEM), Neuchâtel, Switzerland. [3]cMACS, Department of Microbial and Molecular Systems, KU Leuven, 3001 Leuven, Belgium. [4]Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, QLD 4072, Australia. [5]School of Mathematics and Physics, The University of Queensland, Brisbane, QLD 4072, Australia. [6]Advanced Instrumentation for Nano-Analytics (AINA), Luxembourg Institute of Science and Technology (LIST), Materials Research and Technology Department, 41, Rue du Brill, L-4422 Belvaux, Luxembourg. [7]University of Luxembourg, 2 Avenue de l'Université, L-4365 Esch-sur-Alzette, Luxembourg. [8]Institute of Chemical Sciences and Engineering (ISIC), École Polytechnique Fédérale de Lausanne, Sion, Switzerland.
*Corresponding author. Email: xinyu.chin@csem.ch

optical $E_g$ of ~1.70 eV was selected (fig. S2). Some of these stacks were then completed by a $C_{60}$ electron-transport layer (ETL) thermally evaporated in high-vacuum conditions.

The QFLS of bare perovskite layers deposited on glass (left column) remains constant when incorporating FBPAc (Fig. 1A). When inserting the HTL (Fig. 1A, hollow diamond), the addition of FBPAc increases the QFLS slightly with all three HTLs (+20 meV). Overall, QFLS losses relative to the glass/perovskite reference are highest on spiro-TTB (−80 and −45 meV compared with the glass/perovskite reference, without and with FBPAc, respectively), followed by TaTm (−65 and −40 meV), and lowest on Me-4PACz (−25 and −10 meV). These differences suggest improved surface passivation with Me-4PACz (*9*). Without FBPAc, the addition of the $C_{60}$ ETL (Fig. 1A, filled diamond) on the perovskite leads to substantial losses in QFLS (−100 to −50 meV in addition to the HTL-induced losses) caused by the formation of trap states in the first monolayer of $C_{60}$ (*22*). On the other hand, the incorporation of FBPAc in the perovskite film synthesis mitigates some of these ETL-induced losses, resulting in a +100-meV QFLS gain compared with FBPAc-free films.

Overall, FBPAc suppresses $Pb^{2+}/Pb^0$ surface defects, as shown by x-ray photoelectron spectroscopy (XPS) in Fig. 1B and figs. S3 to S5. The phosphonic group of FBPAc coordinates with Pb-related defects as also demonstrated elsewhere (*20*, *23*, *24*), effectively lowering the population of top-surface nonradiative recombination states (*25*). Furthermore, the positive impact of FBPAc is mainly observed when $C_{60}$ is present: FBPAc likely separates the perovskite from the $C_{60}$ layer, reducing the population of deep trap states forming in $C_{60}$ when it is in direct contact with the perovskite (*11*, *22*). It is important to note that some effects may affect the QFLS data presented in this work. First, the presence of oxygen and humidity during the annealing step and subsequent photoluminescence measurements may lead to a passivation of Pb-related defects (*26*). Second, x-ray diffraction (XRD) patterns show that the addition of FBPAc decreases the fraction of $PbI_2$ crystalline phases remaining at the end of the crystallization process (Fig. 1C and fig. S8)—an effect that should lower the photoluminescence quantum yield and hence the QFLS of films featuring the additive (*27*). Other effects are less likely to explain the QFLS trends we observed. For example, FBPAc and $C_{60}$ do not appear to chemically interact based on additional XPS data (fig. S6). Also, the electronic band structure of uncoated perovskite films treated with the additive shifts by ~0.5 eV because of the strong dipole moment of FBPAc (fig. S7), in line with literature (*28*). However, FBPAc is unlikely to lead to such a conduction band level offset in full devices considering the improved solar cell performance achieved with FBPAc (see below).

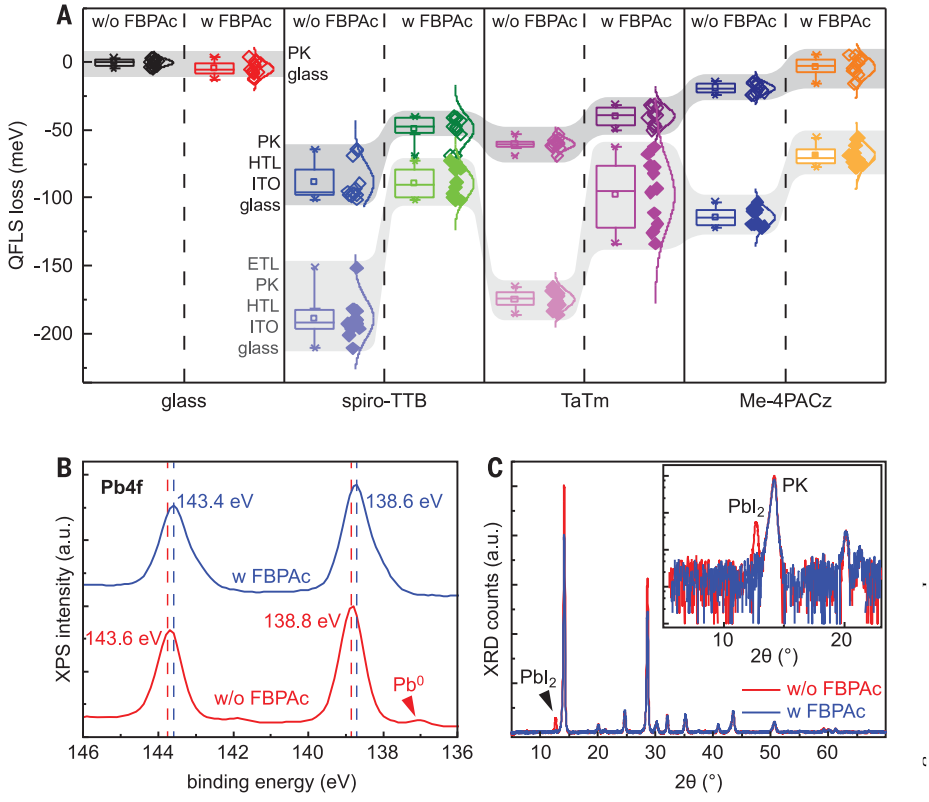



**Fig. 1. Reduced nonradiative recombination losses and improved crystallographic properties when adding FBPAc.** (**A**) QFLS loss of perovskite films (without and with FBPAc) extracted from photoluminescence quantum yield data relative to the glass/perovskite reference. Hollow diamond data points correspond to the QFLS of perovskite films deposited on different HTLs (spiro-TTB, TaTm, and Me-4PACz), and filled diamonds relate to films sandwiched by one of these HTLs and a $C_{60}$ ETL. (**B** and **C**) High-resolution XPS spectra of the Pb 4f orbital (B) and XRD patterns of perovskite films with and without FBPAc (C). a.u., arbitrary units.

## Modification of the perovskite crystallization kinetics and resulting microstructure

To understand the impact of FBPAc on the perovskite crystallization process, grazing-incidence wide-angle x-ray scattering (GIWAXS) was performed in situ at 150°C starting from a $CsBr/PbI_2$ template freshly spin-coated with FAI/FABr (Fig. 2A and fig. S8). Although the crystallization onset of the perovskite (110) and (002) planes ($q \sim 10.1\ nm^{-1}$) (*29*) is similar with or without the additive, FBPAc slows down the perovskite crystallization rate 100 s into the experiment (Fig. 2A). The intensity of these perovskite reflections then continues to increase throughout the experiment up to ~400 s—an effect not observed without the additive. Without FBPAc, the peak intensity stabilizes after 150 s and remains at a lower value. This modification of the crystallization pathway by FBPAc is also observed when analyzing $PbI_2$ peaks (Fig. 2A). As in Fig. 1C, their evolution confirms a more-complete conversion of the $CsBr/PbI_2$ template to a perovskite phase when using FBPAc. Overall, this two-step and globally slower crystallization kinetics with the additive likely originates from the interaction between $Pb^{2+}$ and FBPAc (*20*, *23*, *24*). FBPAc slows down the conversion of the $CsBr/PbI_2$ template to the perovskite phase because of competition between FBPAc and other perovskite precursors (namely FAI and FABr) for a coordination with $Pb^{2+}$ (*25*, *30*). Advantageously, this difference in formation kinetics results in films with apparently larger and more clearly defined perovskite domains because FBPAc may inhibit the nucleation of new perovskite grains or domains, as shown by top-view scanning electron microscopy (SEM) images (Fig. 2B).

Submicrometer dark regions are observed on perovskite films featuring FBPAc (Fig. 2B). From gas-field ion-source helium and neon ion microscopy coupled with secondary-ion mass spectrometry (SIMS) (Fig. 2C and fig. S9) (*31*), these clusters are found to be rich in F and C, which is indicative of the clustering of the FBPAc excess on the perovskite surface. Films processed without FBPAc do not show any clustering and feature a lower amount of F and C on their surface (Fig. 2, C and D, and

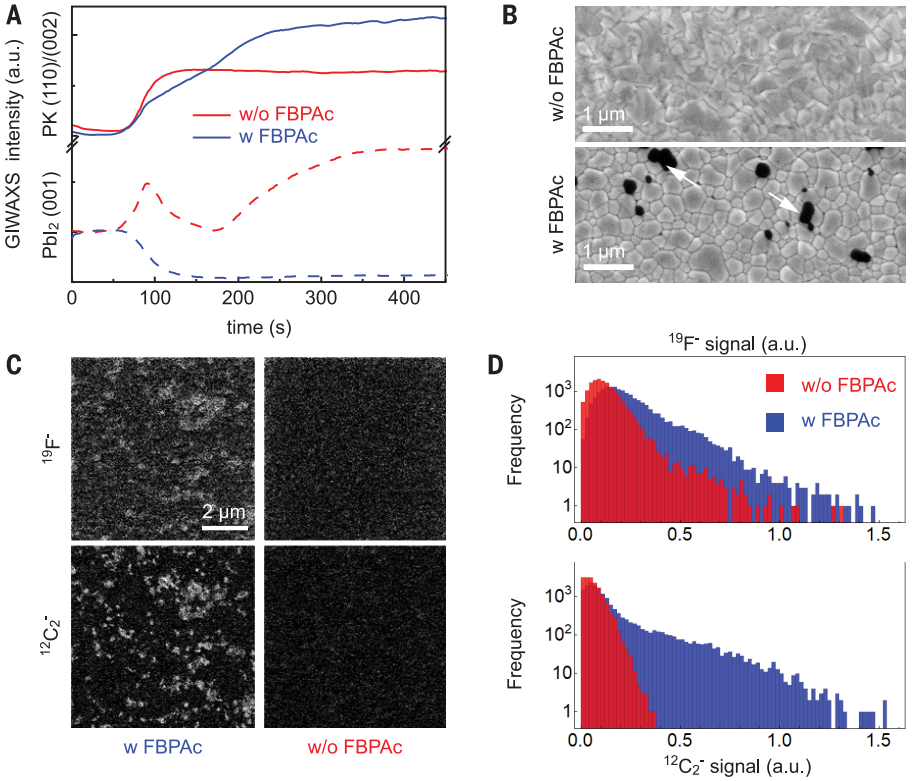



**Fig. 2. Impact of FBPAc on the perovskite crystallization and resulting microstructure.** (**A**) GIWAXS data showing the evolutions of the (110) and (002) planes of the perovskite phase (indicated as PK in the figure) and of the (001) planes of the $PbI_2$ phase during the perovskite crystallization process at 150°C. (**B**) Secondary-electron SEM images highlighting the difference in surface morphologies when adding FBPAc, with arrows showing FBPAc clusters. (**C**) SIMS maps of the $^{19}F^-$ and $^{12}C_2^-$ signals. The bright and dark regions represent higher and lower, respectively, signal counts. (**D**) Corresponding histograms of the $^{19}F^-$ and $^{12}C_2^-$ SIMS signals shown in (C).

fig. S9). Depth profiling combined with XPS reveals that FBPAc is confined to the top perovskite surface (figs. S10 to S13). The bulky FBPAc molecules are expelled toward the top surface during the perovskite crystallization process.

The large reduction in QFLS losses suggests that FBPAc is also present in between the surface clusters. This hypothesis stems from the logarithmic dependency between QFLS and surface recombination current (and hence non-passivated surface fraction, presuming that FBPAc is reducing the surface recombination current by a few orders of magnitude). If FBPAc was present only at the position of the clusters detected by SIMS, meaning that ~90% of the film surface would be free of FBPAc, the difference in QFLS would be limited to ~3 meV [$\Delta QFLS = 25.7\ meV \times \ln(1/0.9) \cong 3\ meV$]. We observe a gain in QFLS of >50 meV when $C_{60}$ is present (Fig. 1A), highlighting that regions between the clusters also feature FBPAc. Overall, FBPAc passivates Pb-related defects and mitigates the formation of trap states in the $C_{60}$ layer by segregating on the perovskite top surface, with the excess forming clusters on the surface.

### Perovskite single junctions and perovskite/c-Si tandem solar cells

These perovskite films were incorporated into ~0.49-$cm^2$ glass/ITO/HTL/perovskite/$C_{60}$/bathocuproine (BCP) (or $SnO_x$)/Cu (or Ag) single junctions to validate the improvements induced by FBPAc with current-voltage (*J-V*) measurements (1000 $W/m^2$, AM1.5G spectrum, masked with a 0.25-$cm^2$ aperture; fig. S14). The open-circuit voltage ($V_{oc}$) of the devices improves by 20 to 30 mV when switching from spiro-TTB to TaTm and Me-4PACz and by an extra 40 to 50 mV when adding FBPAc, in line with the QFLS results of Fig. 1A. The corresponding fill factor (*FF*) also increases on average from ~69 to ~76% when changing the HTL and then to ~78 to 81% with the additive, confirming that the molecule does not lead to the formation of a barrier impeding electron extraction. Overall, the combination of Me-4PACz and FBPAc results in a PCE of up to 19.5%. Varying the concentration of FBPAc from 2.5 to 7.5 mM in the FABr:FAI solution does not affect the performance substantially (fig. S15), which suggests that the excess of FBPAc (Fig. 2B) does not hinder charge transport across the perovskite/$C_{60}$ interface. Washing away the FBPAc with an orthogonal solvent mixture results in perovskite cells with a similar $V_{oc}$ to that of untreated films, which confirms that FBPAc affects primarily recombination at the perovskite top surface (fig. S16).

We then adapted the perovskite deposition conditions to produce 1-$cm^2$ tandem cells on silicon heterojunction bottom cells that had pyramids of 2 to 3 μm on both sides of the wafer (Fig. 3A) (*32*, *33*). Compared with our previous tandem baseline (*16*), the perovskite layer was made thicker (from ~650 nm to ~1 μm) to reach current matching conditions with an $E_g$ of ~1.7 eV (1.63 eV previously; Fig. 3B). For that purpose, we deposited a CsBr/$PbI_2$ template of 660 nm and adapted the spin-coating parameters to convert the template into a perovskite phase. The front electrode was made transparent by switching to a stack made of $C_{60}$, $SnO_x$, indium zinc oxide (IZO), an Ag grid, and a $MgF_2$ antireflective coating (see materials and methods and supplementary text for details).

A $V_{oc}$ of 1.9 V is obtained with either TaTM or Me-4PACz (fig. S17), which is consistent with single-junction cell results (perovskite top cells in fig. S14; c-Si bottom cell in fig. S18). The *FF* of TaTm-based tandems is lower than with Me-4PACz (<70% compared with >75%), likely because of the low hole mobility in undoped TaTm (*34*). We also investigated the impact of a slight change in the perovskite $E_g$ by changing the FABr:FAI molar ratio from 2 to 1.75 and 1.5 (corresponding to a $E_g$ of 1.69, 1.68, and 1.67 eV, respectively). A PCE >30% could be achieved with several of these perovskite compositions, with slight variations in $J_{sc}$ and *FF* values depending on the top cell $E_g$ and the current-limiting subcell (figs. S17 and S19). These tandems remained stable for >360 days when stored at room temperature in $N_2$ (fig. S20). A selection of these unencapsulated cells were then operated in ambient at their maximum power point at 65°C under a 1-sun illumination (~20 to 30% relative humidity; ISOS-L2 protocol). The most stable device reaches a $t_{80}$ of ~66 hours until 80% of its initial PCE is measured (figs. S21 and S22). The degradation rate at 25°C of the FBPAc-treated tandem stack is similar to that achieved without FBPAc in our previous study (ISOS-L1 conditions; figs. S21 and S22) (*16*).

A tandem cell featuring Me-4PACz and FBPAc was sent to the National Renewable Energy Laboratory (NREL) for certification. The certified relative external quantum efficiency (EQE) shown in Fig. 3C indicates that the tandem was slightly top cell limited and that both subcells produced a $J_{sc} > 20\ mA/cm^2$. The stabilized *J-V* properties of the device were measured using the asymptotic maximum power ($P_{max}$) scan method in standard test conditions (STC) (25°C, 1000 $W/m^2$, AM1.5G

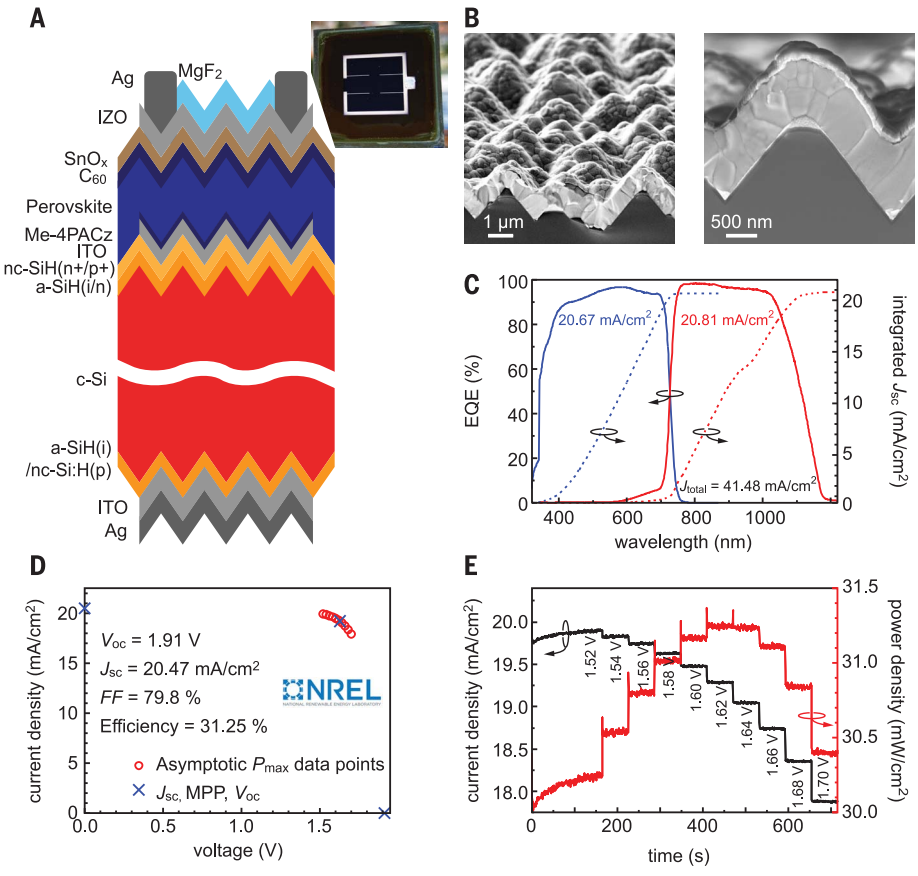



**Fig. 3. 31.25%-efficient perovskite/c-Si thanks to the combined action of Me-4PACz and FBPAc and a pyramidal texture on the front side of the tandem.** (**A**) Schematic view of the perovskite/c-Si tandem stack and corresponding picture of the perovskite top cell on textured Si. The acronyms a-Si:H and nc-Si:H stand for amorphous and nanocrystalline hydrogenated Si, respectively, and (i/n/p) indicates the doping of the layers. (**B**) Secondary-electron SEM images of the front side of the tandem. (**C** to **E**) Independently certified EQE (C), asymptotic maximum power ($P_{max}$) scan (D), and current density and power density tracking around $P_{max}$ (E).

spectrum; Fig. 3, D and E, and figs. S23 to S26) (*35*). In agreement with data measured in-house, the cell reaches a $V_{oc}$ of 1.91 V, a $J_{sc}$ of 20.47 mA/cm$^2$, and a *FF* of 79.8%, which results in a certified PCE of 31.25% (with an aperture area of 1.1677 cm$^2$).

## Conclusions

We identified and mitigated nonradiative recombination losses occurring at the interfaces of perovskite/c-Si tandems featuring Si wafers with a micrometric texture—the standard used in the c-Si PV industry. The use of Me-4PACz reduces voltage losses at the perovskite/HTL interface, whereas the inclusion of FBPAc in the perovskite deposition sequence reduces the voltage losses at the perovskite/$C_{60}$ ETL interface and leads to more favorable perovskite microstructures with larger domains. From XPS and SIMS imaging, FBPAc is present on the perovskite top surface and coordinates with perovskite Pb-related defects through its phosphonic acid group. Overall, the combination of a c-Si micrometric texture to improve optical properties, a high-quality 1-µm-thick perovskite absorber deposited conformally on this texture using a hybrid two-step method, and phosphonic groups at both sides of the absorber for improved interface passivation resulted in a tandem with an independently certified PCE of 31.25%. These results demonstrate how c-Si solar cells with a standard industrial micrometric texture can be upgraded to increase their PCE to >30%.

## REFERENCES AND NOTES


1. "Best Research-Cell Efficiency Chart" (NREL, 2023); https://www.nrel.gov/pv/cell-efficiency.html.
2. S. Schäfer, R. Brendel, *IEEE J. Photovolt.* **8**, 1156–1158 (2018).
3. L. C. Hirst, N. J. Ekins-Daukes, *Prog. Photovolt. Res. Appl.* **19**, 286–293 (2011).
4. F. Fu *et al.*, *Adv. Mater.* **34**, e2106540 (2022).
5. S. De Wolf *et al.*, *J. Phys. Chem. Lett.* **5**, 1035–1039 (2014).
6. X. Y. Chin, D. Cortecchia, J. Yin, A. Bruno, C. Soci, *Nat. Commun.* **6**, 7383 (2015).
7. S. D. Stranks *et al.*, *Science* **342**, 341–344 (2013).
8. T. Jesper Jacobsson *et al.*, *Energy Environ. Sci.* **9**, 1706–1724 (2016).
9. A. Al-Ashouri *et al.*, *Science* **370**, 1300–1309 (2020).
10. B. Chen *et al.*, *Joule* **4**, 850–864 (2020).
11. J. Liu *et al.*, *Science* **377**, 302–306 (2022).
12. P. Tockhorn *et al.*, *Nat. Nanotechnol.* **17**, 1214–1221 (2022).
13. Y. Hou *et al.*, *Science* **367**, 1135–1140 (2020).
14. F. Gota, R. Schmager, A. Farag, U. W. Paetzold, *Opt. Express* **30**, 14172–14188 (2022).
15. D. A. Jacobs *et al.*, *J. Phys. Chem. Lett.* **10**, 3159–3170 (2019).
16. F. Sahli *et al.*, *Nat. Mater.* **17**, 820–826 (2018).
17. C. M. Wolff, P. Caprioglio, M. Stolterfoht, D. Neher, *Adv. Mater.* **31**, e1902762 (2019).
18. J. Peng *et al.*, *Adv. Energy Mater.* **8**, 1801208 (2018).
19. R. Azmi *et al.*, *Science* **376**, 73–77 (2022).
20. E. Akman, A. E. Shalan, F. Sadegh, S. Akin, *ChemSusChem* **14**, 1176–1183 (2021).
21. M. Roß *et al.*, *Adv. Energy Mater.* **11**, 2101460 (2021).
22. J. Warby *et al.*, *Adv. Energy Mater.* **12**, 2103567 (2022).
23. A. A. M. Brown *et al.*, *Nanoscale* **11**, 12370–12380 (2019).
24. J. S. Kim *et al.*, *Nature* **611**, 688–694 (2022).
25. T. Wang, J. D. A. Ng, E. W. Y. Ong, K. H. Khoo, Z. K. Tan, *Adv. Opt. Mater.* **10**, 2101914 (2021).
26. J. F. Galisteo-López, M. Anaya, M. E. Calvo, H. Míguez, *J. Phys. Chem. Lett.* **6**, 2200–2205 (2015).
27. J. Euvrard, O. Gunawan, D. B. Mitzi, *Adv. Energy Mater.* **9**, 1902706 (2019).
28. I. Lange *et al.*, *Adv. Funct. Mater.* **24**, 7014–7024 (2014).
29. L. Oesinghaus *et al.*, *Adv. Mater. Interfaces* **3**, 1600403 (2016).
30. T. Jeon *et al.*, *Adv. Energy Mater.* **7**, 1602596 (2017).
31. T. Wirtz, O. De Castro, J.-N. Audinot, P. Philipp, *Annu. Rev. Anal. Chem.* **12**, 523–543 (2019).
32. F. Sahli *et al.*, *Adv. Energy Mater.* **8**, 1701609 (2018).
33. L. Antognini *et al.*, *IEEE J. Photovolt.* **11**, 944–956 (2021).
34. C. Momblona *et al.*, *Energy Environ. Sci.* **9**, 3456–3463 (2016).
35. T. Song, D. J. Friedman, N. Kopidakis, *Adv. Energy Mater.* **11**, 2100728 (2021).


## ACKNOWLEDGMENTS


The authors thank P. Wyss and C. Allebé for SHJ wet-chemical processing and V. Paratte, B. Kamino, T. Wirtz, J. Hofkens, and M. B. J. Roeffaers for their valuable scientific and technical discussion. The authors also thank the staff of the BL11 NCD-SWEET beamline at ALBA Synchrotron for their assistance in recording the GIWAXS data. **Funding:** This study was supported by Swiss National Science Foundation grant 200021_197006 (PAPET); Swiss Federal Office of Energy grant SI/501804 (INTENT); Swiss Federal Office of Energy grant SI/502209 (PRESTO); Luxembourg National Research Fund (FNR) grant PRIDE17/12246511 (PACE); and Research Foundation - Flanders FWO, grant no. 12Y7221N, V400622N. C.M.W. thanks the European Commission for funding through a Marie Skłodowska-Curie Fellowship (no. 101033077). **Author contributions:** Conceptualization: X.Y.C. and Q.J. Experiments: X.Y.C., D.T., J.A.S., S.T., S.E., M.M., P.F., C.M.W., A.P., K.A., D.J., Q.G., F.S., G.A., M.B., and Q.J. Project administration: X.Y.C., C.M.W., G.A., M.B., Q.J., and C.B. Supervision: X.Y.C., C.M.W., Q.J., and C.B. Writing – original draft: X.Y.C. and Q.J. Writing – review & editing: X.Y.C., D.T., C.M.W., J.A.S., S.T., S.E., M.M., P.F., A.P., K.A., D.J., Q.G., F.S., G.A., M.B., Q.J., and C.B. **Competing interests:** X.Y.C. and Q.J. are inventors on a provisional patent related to the subject matter of this manuscript (EP22210577.7). The authors declare no other competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.


**License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg0091
Materials and Methods
Supplementary Text
Figs. S1 to S28
Tables S1 to S4
References (*36*–*45*)


Submitted 27 November 2022; accepted 5 May 2023
10.1126/science.adg0091

SOLAR CELLS

# Interface engineering for high-performance, triple-halide perovskite–silicon tandem solar cells

Silvia Mariotti[1]†‡, Eike Köhnen[1]†, Florian Scheler[1], Kári Sveinbjörnsson[1], Lea Zimmermann[1], Manuel Piot[1,2], Fengjiu Yang[1], Bor Li[1], Jonathan Warby[3], Artem Musiienko[1], Dorothee Menzel[1], Felix Lang[3], Sebastian Keßler[1], Igal Levine[1], Daniele Mantione[4,5,6], Amran Al-Ashouri[1], Marlene S. Härtel[1], Ke Xu[1], Alexandros Cruz[1], Jona Kurpiers[1], Philipp Wagner[1], Hans Köbler[1], Jinzhao Li[1], Artiom Magomedov[7], David Mecerreyes[4,6], Eva Unger[1], Antonio Abate[1], Martin Stolterfoht[3], Bernd Stannowski[1], Rutger Schlatmann[1], Lars Korte[1], Steve Albrecht[1,8]*

Improved stability and efficiency of two-terminal monolithic perovskite-silicon tandem solar cells will require reductions in recombination losses. By combining a triple-halide perovskite (1.68 electron volt bandgap) with a piperazinium iodide interfacial modification, we improved the band alignment, reduced nonradiative recombination losses, and enhanced charge extraction at the electron-selective contact. Solar cells showed open-circuit voltages of up to 1.28 volts in p-i-n single junctions and 2.00 volts in perovskite-silicon tandem solar cells. The tandem cells achieve certified power conversion efficiencies of up to 32.5%.

Although the photovoltaic (PV) devices market is currently dominated by crystalline silicon (c-Si) devices, perovskite-silicon tandem architectures have gained attention to increase their power conversion efficiency (PCE). Two-terminal devices that directly connect a c-Si bottom cell and perovskite top cell use light more efficiently because high-energy photons can be absorbed by the perovskite top cell, and low-energy photons transmitted by the perovskite can be absorbed by the c-Si bottom cell (*1*). This approach recently led to a PCE of 33.7% (*2*). However, given that the detailed balance limit of monolithic perovskite-silicon tandems is 45.1% (*3*), further improvements could be made, especially concerning the open-circuit voltage ($V_{OC}$) (*1*, *4*). To reach the highest possible efficiency under the air mass coefficient (AM) 1.5G solar spectrum and current-matching conditions, the optimum band gap $E_G$ of the perovskite is 1.73 eV (*3*). However, finite perovskite thickness and parasitic absorption reduce the optimal perovskite band gap to ~1.68 eV (*5*–*8*). Perovskite compositions with this band gap require high bromide contents and suffer from instability caused by phase segregation that lead to large $V_{OC}$ deficits (*1*). To address this problem, we combined an efficient perovskite composition with an interface modification. Thorough analysis reveals the mechanism leading to high $V_{OC}$-values and hence, high PCE-values in single-junction solar cells and perovskite-silicon tandem solar cells.

[1]Solar Energy Division, Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, 12489 Berlin, Germany. [2]Laboratory of Semiconductor Materials, Institute of Materials, Faculty of Engineering, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland. [3]Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany. [4]POLYKEY Polymers, Joxe Mari Korta Center, 20018 Donostia-San Sebastian, Spain. [5]POLYMAT, University of the Basque Country UPV/EHU, 20018 Donostia-San Sebastián, Spain. [6]IKERBASQUE, Basque Foundation for Science, 48009 Bilbao, Spain. [7]Kaunas University of Technology, Kaunas, Lithuania. [8]Technische Universität Berlin, Fakultät Elektrotechnik und Informatik, 10587 Berlin, Germany.
***Corresponding author. Email: steve.albrecht@helmholtz-berlin.de**
†These authors contributed equally to this work.
‡Present address: Energy Materials and Surface Sciences Unit (EMSSU), Okinawa Institute of Science and Technology Graduate University (OIST), Okinawa 904-0495, Japan.

## Perovskite compositions and surface treatment

A widely used perovskite composition is the "triple-cation" (3Cat) perovskite $Cs_{0.05}(FA_{0.77}MA_{0.23})_{0.95}Pb(I_{0.77}Br_{0.23})_3$, which contains cesium (Cs), methylammonium (MA), and formamidinium (FA) as cations and mixed bromine (Br) and iodine (I) anions (*7*, *8*). However, the so-called "triple-halide" (3Hal) composition $Cs_{0.22}FA_{0.78}Pb(I_{0.85}Br_{0.15})_3$ + 5% $MAPbCl_3$ substantially reduces the MA content, which is associated with instability, while increasing the amount of Cs and using three halides (adding $Cl^-$ to $I^-$ and $Br^-$) to improve the quality of the perovskite crystal lattice, stability (*9*, *10*), and minimized phase segregation (*11*). In addition, the 3Hal has shown substantial improvement of quasi-Fermi level splitting (QFLS) and a low ideality factor (*12*) compared with those of more common perovskite compositions, which are factors that generally determine high $V_{OC}$ values. By naming this perovskite 3Hal, we do not exclude the presence of three cations, but we emphasize the presence of three halides that confer the positive properties to this perovskite. Previous reports showed that this 3Hal perovskite enables efficient tandem cells with a PCE of 27% (*13*) and has a pseudo-PCE [obtained with photoluminescence (PL) measurements] of 33.4% (versus 32.6% for tandems with 3Cat perovskite) (*12*). Thus, 3Hal was used for this study.

Besides tuning the composition of the perovskite layers to improve performance, additives are particularly effective in passivating electronic traps and eliminating current density–voltage (*JV*) hysteresis in devices (*14*). Chemical passivation of the perovskite layer can reduce nonradiative recombination lo… that occur in the bulk of the film (particul… at the grain boundaries) or at the interfaces with the charge transport layers (CTLs), which can improve device $V_{OC}$ and stability (*15*–*17*). We focused on piperazinium iodide (PI), an ionic liquid previously reported by Li *et al.* (*18*) for use with a FA-rich 3Cat perovskite. The synthesis and characterization of the PI molecule are presented in the supplementary materials (fig. S1). Li *et al.* showed that PI contains both electron acceptor ($R_2NH_2^+$) and electron donor ($R_2NH$) functional groups that can interact with both positive and negative perovskite surface traps caused by $Pb^{2+}$ and halide ions, reducing nonradiative recombination (*18*). In addition to the chemical passivation of the perovskite surface, Li *et al.* show that the formation of a dipole between the perovskite and the electron-transport layer (ETL) participates in reducing $V_{OC}$ losses to 330 mV in p-i-n devices (*18*).

We combined the 3Hal perovskite with a bandgap suitable for tandem integration (1.68 eV) with PI interfacial modification and compared it with the commonly used lithium fluoride (LiF) interlayer (*19*, *20*). As do other alkali-earth metal fluorides, such as $MgF_2$, LiF acts as a contact displacer that improves the device performance by favorably adjusting the surface energy of the perovskite layer in contact with the fullerene-based ETL ($C_{60}$) (*17*). By contrast, PI shows a faster charge extraction and reduced nonradiative recombination losses and creates a more favorable band alignment between the perovskite and the $C_{60}$ ETL by forming a positive dipole, while no chemical passivation of the perovskite surface is detected. This modification allows $V_{OC}$ values of up to 1.28 and 2.00 V and PCE values of 21.5 and 32.5% for perovskite single-junction and perovskite-silicon tandem solar cells, respectively.

## Electronic properties of 3Hal films with surface passivation and charge-transport layers

Several characterization methods were used to determine the film properties of the perovskite layer and the effect of the surface treatment, with and without CTLs. The following three configurations were considered, with either no surface treatment or PI or LiF surface treatment: (i) quartz glass with the perovskite directly deposited onto this substrate; (ii) "half stack" devices (glass/ITO/2PACz/3Hal), where ITO is indium tin oxide and 2PACz is [2-(3,6-dimethoxy-9H-carbazol-9-yl)ethyl]phosphonic acid; and (iii) "full stack" devices (glass/ITO/2PACz/3Hal/$C_{60}$) (*21*).

We performed PL measurements on the three configurations mentioned above to obtain PL quantum yield (PLQY) and quasi-Fermi level splitting (QFLS) values (table S1). The results show that the highest QFLS was obtained for the LiF samples on quartz (1.33 eV) and in the

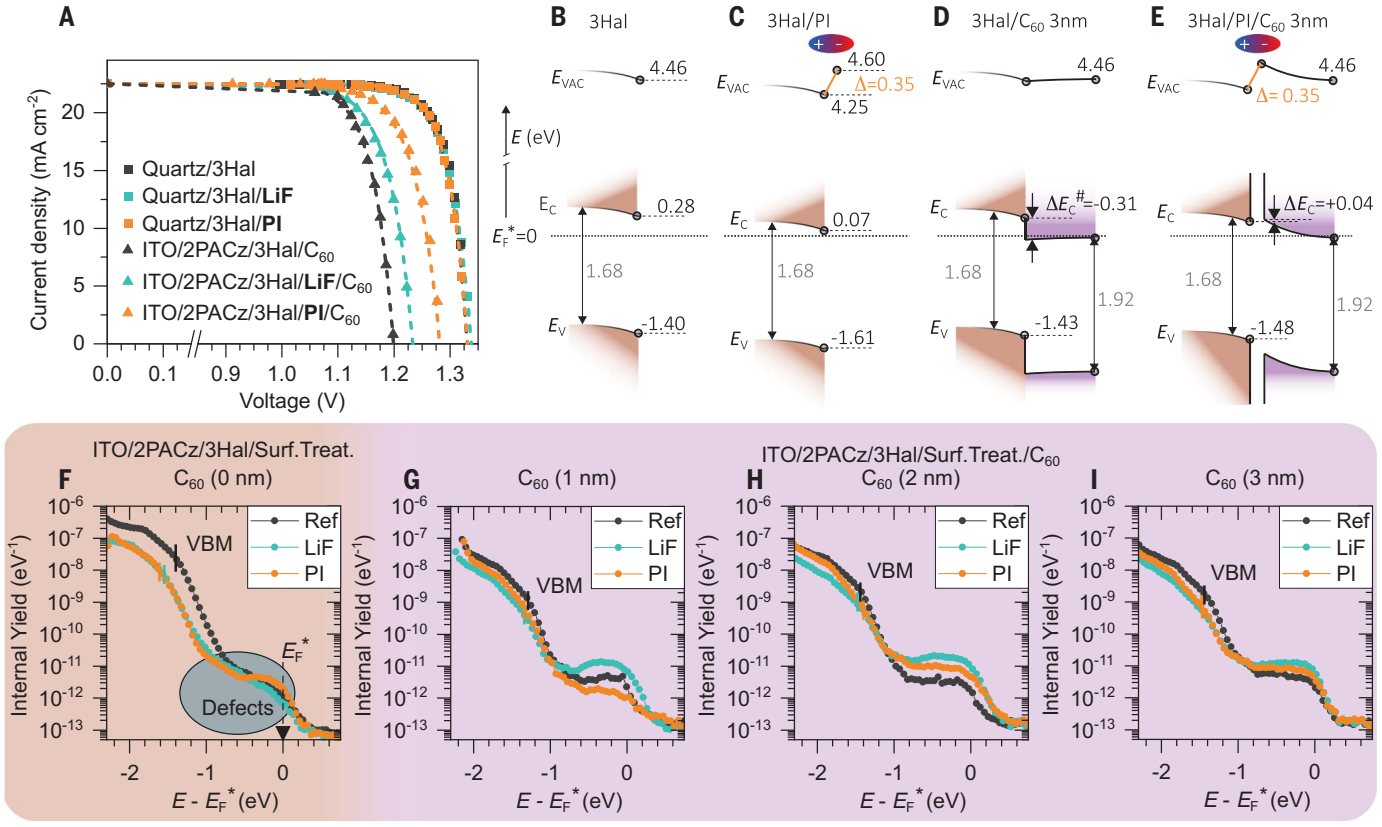



**Fig. 1. Electronic properties of 3Hal perovskite films with surface treatment and charge-transport layers.** (**A**) Pseudo-*JV* curves for samples without surface modification, with LiF and PI, both for single films on quartz and for complete device stacks. (**B** to **E**) Energy-level alignment of the 3Hal-PI and 3Hal-PI-$C_{60}$ interface, obtained by means of UPS ($h\nu$ = 6.5 eV) and CFSYS. The energetic positions of the 3Hal perovskite are obtained from the respective drawn layer stack—for (D) and (E), the perovskite VBM obtained from below the 3 nm $C_{60}$ is used. $\Delta E_C{}^{\#}$ presented in (D) is the conduction band offset taken from (*34*). All values presented in the figure are in electron volts. (**F** to **I**) Valence band region and occupied gap state density of 3Hal perovskite with LiF and PI for varying $C_{60}$ thickness [0, 1, 2, and 3 nm for (F), (G), (H) and (I), respectively] measured with CFSYS on the stack ITO/2PACz/3Hal/Surf.Treat./$C_{60}$. The energy reference zero was obtained by modeling the Fermi edge ($E_F$*) at the sample surface observed in the spectra.

half stack (1.29 eV) configurations, but for the full stack, the PI sample shows the highest QFLS (1.27 eV). From the intensity-dependent PL results, pseudo-*JV* curves were constructed (Fig. 1A). For samples on quartz, LiF achieved the highest PLQY (9%) of the series, whereas both reference and PI samples show similar PLQY (~5%). The negligible impact of PI on the PLQY and pseudo-*JV* of perovskite-on-quartz samples rules out a chemical passivation effect, but interface recombination occurs in the full stack with ITO and the CTLs. Warby *et al.* demonstrated that the recombination is mainly caused by the presence of $C_{60}$ (ETL) (*21*). Although perovskite-on-quartz samples show similar pseudo-*JV* curves and hence similar pseudo-PCE for samples with and without treatment, in full stack configuration we observed a higher pseudo-PCE for PI owing to a higher QFLS. The QFLS of the full-stack at 1 sun equivalent illumination are 1.27, 1.23, and 1.18 eV for samples with PI treatment, with LiF treatment, and without treatment, respectively. The beneficial effect of the PI occurred only with $C_{60}$ as ETL, indicating that PI used on 3Hal perovskite reduces across-interface recombination rather than directly passivating the perovskite surface.

We performed near–ultraviolet (UV) photoelectron spectroscopy (UPS) measurements ($h\nu$ = 6.5 eV, where $h$ is Planck's constant and $\nu$ is the photon frequency) as well as with photon energies varying from 4 to 7 eV in the constant final-state yield mode (CFSYS) (*22*) to investigate the band alignment at the interface of 3Hal-$C_{60}$ with LiF or PI interface modification and $C_{60}$ thickness varying from 0 to 3 nm in the full stack samples. We applied a detailed model of the perovskite valence band maximum (VBM) in combination with the work function (WF) to precisely determine the ionization energy and thus assess possible dipole effects of the interface modification (figs. S2 and S3). The WF is a property of the topmost layer in the investigated stack. For 3Hal-LiF, the WF decreased by ~−0.17 eV compared with the nontreated perovskite, which is consistent with our previous study (*22*). However, the WF of 3Hal-PI increased compared with the reference sample by 0.14 eV, and hence the PI-treated surface shows the highest WF of 4.60 eV (Fig. 1C and table S2) (*23*). Kelvin Probe (KP) revealed larger WFs for samples with PI surface treatment compared with the untreated reference with increase of 0.24 and 0.15 eV with and without 2PACz, respectively (tables S3 and S4).

Changes in the WF of a semiconductor surface can be caused by a shift in the Fermi level position within the band gap (through doping or charge redistribution), by a surface dipole, or both effects. We distinguished these effects by measuring the ionization energy $I$*, the difference between the perovskite VBM $E_V$ and the vacuum level $E_{Vac}$. The formation of a dipole changes both WF and $I$* by the same amount, whereas a change in $E_F$ within the band gap leaves $I$* unchanged. We determined the $E_V$ and thus calculated $I$* not only on the neat, LiF-treated, or PI-treated surfaces, but also underneath thin layers of $C_{60}$ (up to 3 nm), thus also accounting for potential charge redistribution occurring at the interface. Without $C_{60}$, $I$* was 5.86, 5.84, and 6.21 eV for samples without an interlayer and with LiF

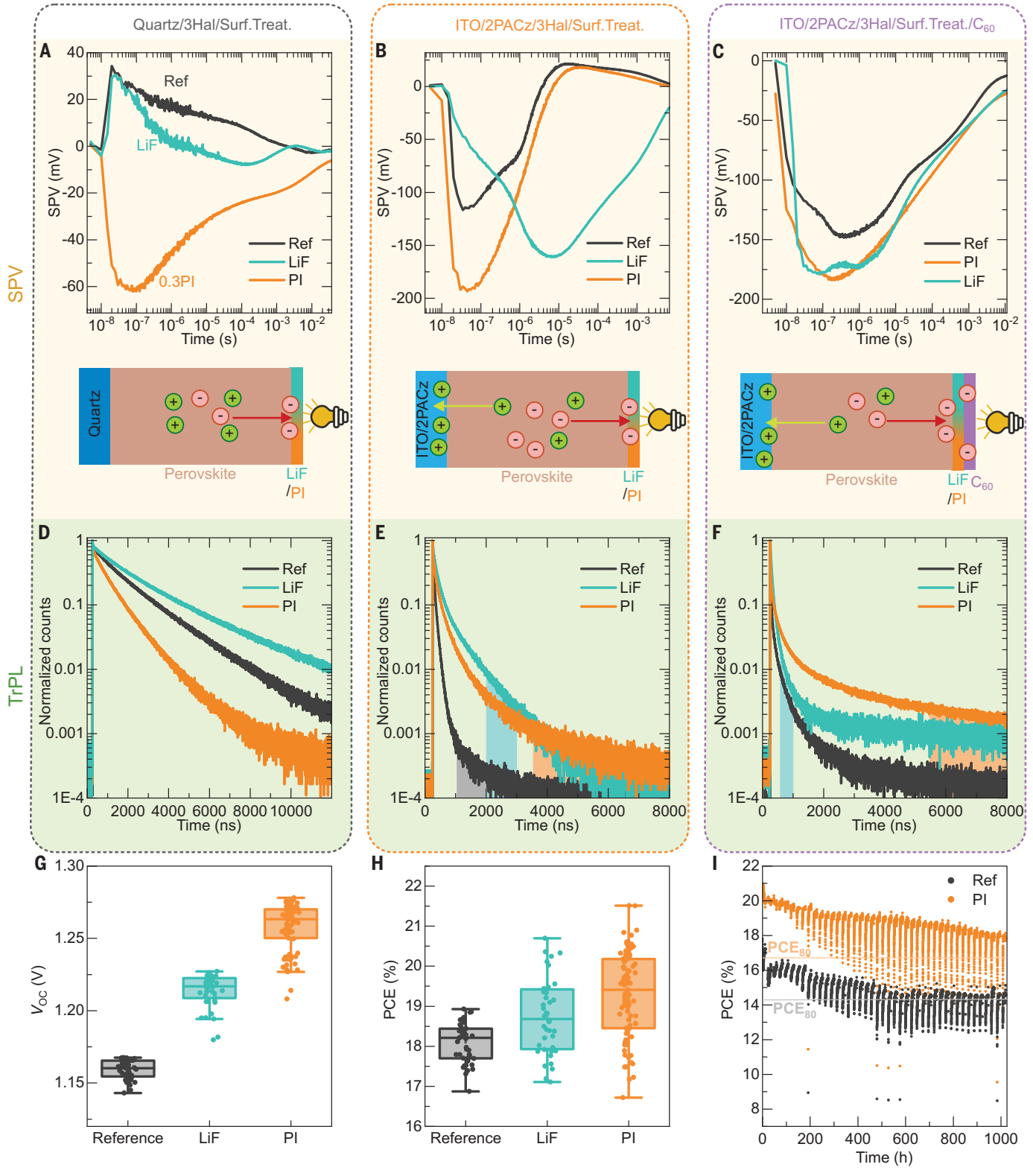



**Fig. 2. Properties of 3Hal films and performance of single-junction solar cells with and without surface treatment.** (**A** to **C**) tr-SPV graphs showing the possible charge separation mechanisms at the bottom, and (**D** to **F**) TRPL analysis performed on different stacks: (A) and (D) quartz/3Hal, (B) and (E) half stack (ITO/2PACz/perovskite), and (C) and (F) full device, that is, with a $C_{60}$ layer added, all considering 3Hal with and without LiF or PI surface treatment. The fluence was ~110 nJ cm$^{-2}$, corresponding to around 6 suns equivalent steady-state charge carrier density. For SPV, signals were measured after a 5-ns laser excitation with a fluence of 38.3 μJ cm$^{-2}$ and photon energy of 1.8 eV. For TRPL, single exponentials were used for fitting the curves. For the half stack and full device cases, the shaded areas indicate the data range in which mono-exponential fits were conducted to approximate the nonradiative decay

time (after the initial decay, before the background noise level). (**G** and **H**) Box charts presenting statistics of $V_{OC}$ and PCE respectively, of the p-i-n single junctions by using 3Hal perovskite with PI and LiF treatment. (**I**) Aging test of perovskite single-junction devices according to ISOS-LC-1l, prepared by using 3Hal as absorber layer with or without PI surface treatment. Devices were MPP-tracked in nitrogen at 25°C for 1020 hours under cycled light (12 hours dark, 12 hours illuminated). Averages of 11 (PI) and eight (reference) solar cells are shown. The inset line corresponds to the 80% of the maximum PCE of the first cycle ($PCE_{80}$).

and PI, respectively. Although LiF might slightly reduce $I^*$, PI increased it by ~350 meV, indicating the presence of an electrostatic dipole at the sample surface, as also observed by Li *et al.* (*18*) although oriented in the opposite direction. Possible reasons for this difference include the different chemical composition of the perovskite they used (3Cat, low bandgap) and the higher amount of unreacted $PbI_2$ on the perovskite surface detected with x-ray diffraction, caused by the loss of $MA^+$ or $FA^+$ cations during degradation processes. The variations were less pronounced with an additional $C_{60}$ layer: the $C_{60}$ thickness-averaged $I^*$ of the perovskite underneath $C_{60}$, LiF-$C_{60}$, and PI-$C_{60}$ were 5.88, 5.88, and 5.95 eV, respectively. Thus, whereas $I^*$ with and without LiF are similar within error, for PI, $I^*$ increased by ~70 meV.

In Fig. 1, B to E, we combine the photoelectron spectroscopy (PES) results for the neat 3Hal and the 3Hal-$C_{60}$ interface with and without PI surface treatment. The sketched band bending in the perovskite is based on the assumption of a more intrinsic bulk (*24*), which the surface-sensitive PES measurements cannot probe. Considering the conduction band offset ($\Delta E_C$) as determined for the 3Cat-$C_{60}$ interface and using the slightly higher band gap energy of 3Hal, we obtained $\Delta E_C$ = –0.31 eV (Fig. 1D). The positive dipole induced by the PI molecule reduced this offset between the 3Hal conduction band minimum and the $C_{60}$ lowest unoccupied molecular orbital (LUMO) to +0.04 eV. We propose that the reduced interfacial losses upon PI incorporation originated from changes in WF induced by a positive dipole (*23*), which shifts all energy levels of the $C_{60}$ at the interface, which in turn decreases the conduction band offset and enables a high $V_{OC}$.

We also used CFSYS measurements to characterize the density of occupied states in the perovskite band gap. On samples without ETL (Fig. 1F, 0 nm $C_{60}$), we observed a similar gap state density for 3Hal and 3Hal-LiF, whereas for 3Hal-PI, it appeared to be slightly increased, especially near the Fermi edge. This finding agrees with PL measurements, which showed that the QFLS could only be increased with PI in the presence of $C_{60}$, indicating the lack of chemical passivation with the PI treatment. For the samples with $C_{60}$ (Fig. 1, G to I, 1 to 3 nm) on the perovskite, the gap states formed at the interface contributed to the CFSYS spectrum because the probing depth is between 5 and 10 nm. The sample with LiF showed the highest gap state density for all three thicknesses (1, 2, and 3 nm), as observed in our previous study (*22*). For 1 nm $C_{60}$, the lowest gap state density was found for the PI-modified interface, which is consistent with PI reducing the nonradiative recombination. Nevertheless, for thicker $C_{60}$ layers, the reference layer stack showed the lowest gap state density. With these and the PL data previously presented, we attribute the positive influence of the PI to the formation of a positive dipole on the 3Hal perovskite.

In solar cells, the extraction of photogenerated electrons and holes occurs simultaneously; hence, it is challenging to separate both processes. We measured transient surface photovoltage (trSPV), which is proportional to the concentration of separated charge carriers at the respective CTLs, to study charge extraction, which appears as a fast initial rise in the SPV signal, and recombination, which appears as the decay of the SPV. Poor selectivity, slow carrier extraction, and severe trap-assisted recombination also slow down the initial increase of the SPV signal and decrease the SPV amplitude (*25*, *26*). Thus, we focused on the previous configurations—namely, perovskite films on quartz glass without CTLs, half stack and full stack devices—to disentangle the effect of passivation on electron and hole extraction dynamics and recombination. The effect of surface treatment on perovskite films on quartz glass without CTLs is shown in Fig. 2A. The PI-treated structure exhibited a substantial negative amplitude, corresponding to the flow of electrons to the perovskite surface and free holes prevailing in the bulk of the perovskite film. The excess electron concentration near the surface was consistent with a ~200-meV-lower perovskite VBM for the PI-treated surface found from PES measurements, suggesting a higher n-type character of the surface (Fig. 1, B and C), as also reported by Li *et al.* (*18*).

We found that the SPV signal for the PI sample depends on its surface coverage and that the maximum negative amplitude was observed for a 0.3 mg $ml^{-1}$ PI solution (fig. S4). The perovskite without any surface treatment showed a positive SPV signal revealing hole accumulation or trapping near the surface (ETL side of the structure). By contrast, the LiF treatment decreased the positive amplitude, implying passivation of the perovskite surface and a reduced concentration of trapped or accumulated holes near the ETL surface.

To investigate hole extraction, we studied the hole-transport layer (HTL)–perovskite interface in the half stack configuration (Fig. 2B). The SPV signal shows an almost doubled amplitude and a steeper slope (increased extraction rate) for the PI-treated samples compared with the nontreated perovskite samples, which we attributed to reduced recombination and improved electron charge selectivity at the ETL side (*26*), as observed for samples without any CTL (Fig. 2A). However, the LiF samples with HTL showed a slower or delayed rise in the SPV signal as compared with that of reference and PI samples. Similar responses were observed in the full stack configuration presented in Fig. 2C, in which free electrons and holes are extracted at the same time, into their respective transport layers. The PI-treated sample demonstrated the fastest rise and largest amplitude, whereas LiF and reference samples showed a slower rise and lower amplitude. Delayed charge extraction could be caused by a tunneling barrier to electron extraction introduced by LiF (*17*), which is avoided when using PI.

Additionally, the time-resolved PL (TRPL) analysis in Fig. 2, D to F, showed a fast extraction process of about 10 ns that is followed by long tails (in the order of a few 100 ns to a few microseconds) that are governed by nonradiative recombination (*7*). The decay times extracted by mono-exponential fits to the transient at later times (at ~2 to 5 μs) allowed us to assess the nonradiative losses. 3Hal perovskite on quartz (Fig. 2D) showed a carrier lifetime ($\tau_1$) of 2 μs, which is in agreement with measurements on a 3Hal perovskite layer in our recent study (*27*). Compared with perovskite on quartz, a reduced $\tau_1$ was observed when adding 2PACz to the stack (from 2 to 0.4 μs for the bare perovskite on quartz and half stack, respectively). In the half stack (Fig. 2E), the perovskite with PI showed a decay time around twice as large as with a bare perovskite (~0.7 μs), which indicated that PI slightly reduced nonradiative recombination losses at the perovskite surface. When $C_{60}$ was added to the stack (Fig. 2F), the recombination losses increased because of the high gap state density and the energy level offset (0.31 eV) present at the perovskite-$C_{60}$ interface, further reducing $\tau_1$ from ~400 to 200 ns, considering the bare perovskite only. However, with PI, the decay time increased drastically ($\tau_1$ ~ 4 μs) and was about five times longer than without $C_{60}$ (0.7 μs). For LiF surface treatment on the full stack device, $\tau_1$ does not change compared with the bare perovskite (both show $\tau_1$ ~ 0.2 μs). Thus, the TRPL transients suggest that PI at the concentration of 0.3 mg $ml^{-1}$ helps retain the high PL of 3Hal when paired with $C_{60}$, as shown by the enhanced carrier lifetime.

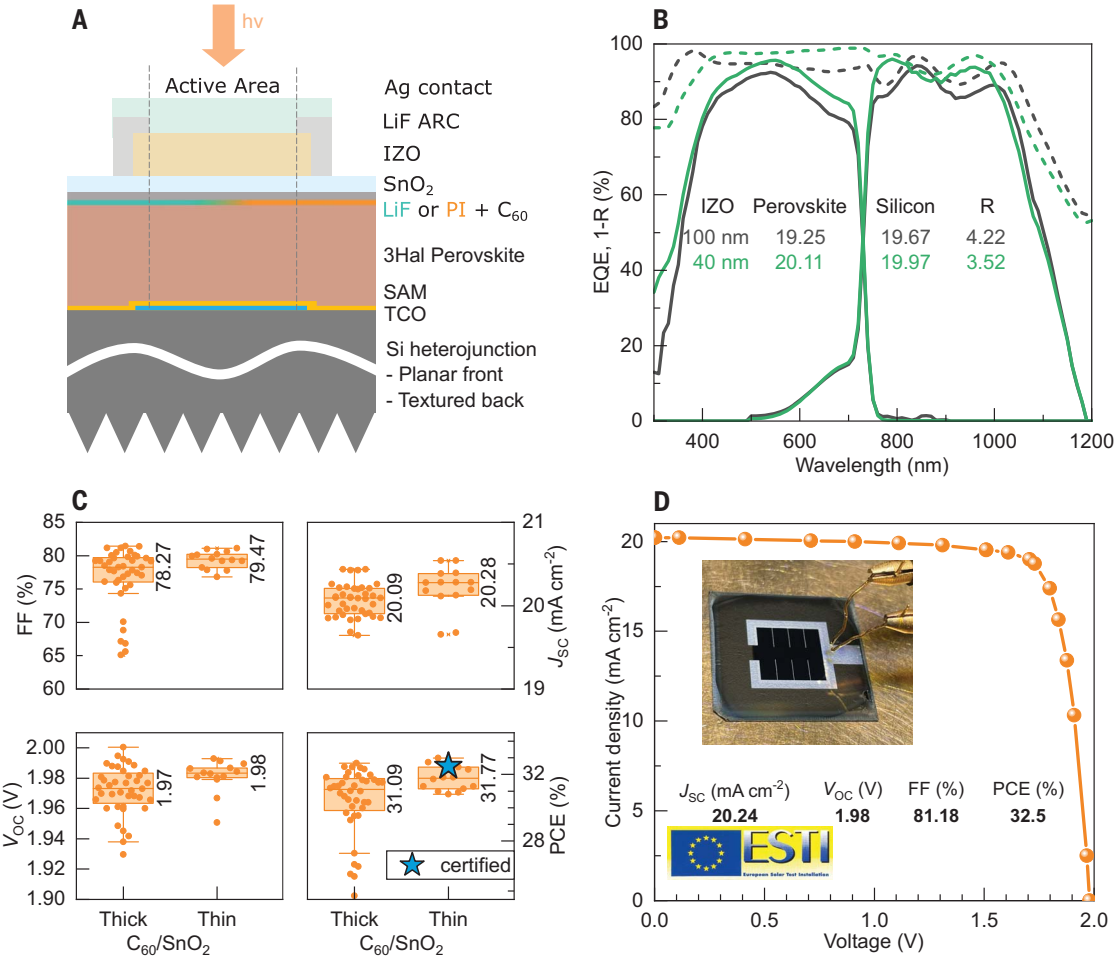



**Fig. 3. Performance of optimized perovskite-silicon tandem solar cells.** (**A**) Sketch of the perovskite-silicon tandem device stack with a silicon (Si) heterojunction bottom cell, transparent conductive oxide (TCO), self-assembled monolayer (SAM), triple-halide (3Hal) perovskite, lithium fluoride (LiF) or piperazinium iodide (PI), $C_{60}$, tin oxide ($SnO_2$), zinc-doped indium oxide (IZO), LiF antireflective coating (ARC), and a silver (Ag) electrode. The arrow indicates the illumination (denoted as *hv*). (**B**) Measured EQE and reflection (denoted as 1-R) of tandem solar cells with 100 and 40 nm IZO. The cell with 40 nm IZO has grid fingers on top and therefore includes shading because the EQE is measured with a spot larger than the grid spacing. The given numbers are the integrated current densities (under AM1.5G illumination) for the reflection (R) and both subcells, in units of milliamperes per square centimeter. (**C**) In-house measured performance of tandem solar cells after implementing PI, Me-4PACz, RDBL, 40 nm IZO, grid fingers and thick (18 + 20 nm) or thin (16 + 10 nm) $C_{60}$/$SnO_2$. There were 20 devices for thick and seven devices for thin $C_{60}$/$SnO_2$. The horizontal line in the boxes and the value denotes the median value. The blue star indicates the values certified at JRC-ESTI. This point does not contribute to the statistical box because the cell is already included in the in-house measurements. (**D**) Certified (quasi-) steady-state *JV* measurement conducted at JRC-ESTI, including the performance parameters and a photograph of the tandem cell.

### Structural properties of 3Hal films with surface passivation and charge-transport layers

Scanning electron microscopy (SEM) top view images of the perovskite films are shown in fig. S5. The 3Hal grains (200 to 300 nm wide) showed similar morphologies with and without PI, which indicated that PI formed either a film that is too thin to be detected or that it did not create a uniform layer on the perovskite surface. After LiF evaporation, agglomerates, likely LiF clusters, formed on the perovskite surface rather than forming a uniform layer (fig. S5D). In contrast to LiF, which has been shown to act as a contact displacer between the perovskite and the $C_{60}$ layer (*17*), for PI the presence of a layer on the surface is uncertain: Energy dispersive x-ray analysis (EDX) and x-ray photoelectron spectroscopy (XPS) showed a slight increment of the nitrogen (N), carbon (C), and hydrogen (H) signals and N–C bond, respectively, but this is not sufficient to determine the presence of a PI layer and/or PI agglomerates (figs. S6 and S7). We performed grazing-incidence wide-angle x-ray scattering (GIWAXS) (fig. S8) on perovskite films with and without surface treatment (both PI and LiF). Although the perovskite (100) pattern increases slightly for the PI-treated perovskite, the measurement reveals that post-deposition treatments did not influence grain growth. Therefore, we assume that PI only plays a role in improving the band alignment rather than acting as contact displacer.

### Single-junction solar cells

We fabricated single-junction devices using the following device stack: Glass/ITO/2PACz (SAM)/3Hal/[surface treatment]/$C_{60}$/tin oxide ($SnO_2$)/silver (Ag). Details regarding the device preparation are presented in the supplementary materials. Shown in Fig. 2, G and H, are the single-junction $V_{OC}$ and PCE values by using 3Hal as absorbing material with and without PI and LiF (PI concentration optimization is shown in fig. S9, with an optimum at 0.3 mg $ml^{-1}$, and all device parameters are shown in fig. S10). Surface treatment of 3Hal with PI improved the $V_{OC}$ (median of 1.26 V) compared with LiF (median of 1.22 V). The high $V_{OC}$ values of up to 1.28 V (91.5% of the detailed-balance limit) obtained by using PI agree with the QFLS (Fig. 1A and table S1) and results in a $V_{OC}$ deficit ($E_G$/q-$V_{OC}$) of just 400 mV. This may be explained by the modification of the WF (from UPS and KP) by a dipole, as discussed above and shown in Fig. 1E, fig. S3, and tables S2 and S3. However, the light soaking effect shows almost no $V_{OC}$ decay after ~2 min of light exposure (fig. S11).

The operational stability of devices was assessed in a light-cycling test (12 hours of illumination followed by 12 hours of dark) according to ISOS-LC-1I (*28*) at 25°C in nitrogen for 1020 hours (lamp spectrum is available in fig. S12) (*29*). These measurements can be considered more realistic than a constant illumination test because day-night light cycles also occur under real operation conditions outdoors. The light-cycled test is additionally able

to account for metastable behavior of PSCs. Maximum power point (MPP) tracks (average of 11 and 8 solar cells for PI and reference, respectively) and 80% of the maximum PCE of the respective first cycle ($PCE_{80}$) are shown in Fig. 2I. The PI device shows a maximum value of 20.98% PCE in the first cycle and 17.84% in the last one. Reference devices show a maximum value of 17.81% PCE in the first cycle and 14.58% in the last cycle. Thus, PI devices retain 85.03% and reference devices retain 81.91% of the initial maximum PCE. The evolution of the cycled behavior is shown in fig. S13, where for each cycle a different time to plateau is observed and is dependent on the type of device. PI devices show a faster rise of PCE for all considered cycles. A faster transient behavior is desirable because devices reach their maximum efficiency faster, and more energy is produced within a light cycle. The slow transient behavior in the beginning of a light cycle has shown to be caused by the slow migration of cations (*30*).

### Perovskite-silicon tandem solar cells

A schematic of the tandem configuration stack is shown in Fig. 3A, and the detailed procedure steps are presented in the supplementary materials.

Tandem devices with PI could confirm higher maximum $V_{OC}$ values (1.95 V) compared with that of LiF (1.90 V) (fig. S14). A continuously MPP-tracked (in air at 23° to 32°C) tandem cell with PI retained 80% of its initial PCE (28.2%) after 347 hours and 75.7% after 478 hours (fig. S15).

We implemented several optimizations separately and combined all of them to improve performance. First, the device fabrication was optimized. The self-assembled monolayer (SAM) Me-4PACz {[4-(3,6-dimethyl-9H-carbazol-9-yl) butyl]phosphonic acid} was used as HTL for the top cell, together with a wetting-improving additive (*7*, *31*). In contrast to our previous work (*7*), in this study this HTL mainly improved median $V_{OC}$ values from ~1.92 V (2PACz) to 1.94 V (Me-4PACz) (fig. S16), agreeing with QFLS measurements (table S1). We also increased the photogenerated current density of the silicon bottom cell ($J_{Si}$) by adding a reflector with a dielectric buffer layer (RDBL) on the rear of the bottom cell, as previously implemented by Cruz *et al.* on silicon single-junction (*32*) and tandem solar cells (*8*). This set of modifications enabled a PCE of 29.9%, as presented in fig. S17 together with the MPP tracking, which shows stable values over 3 min.

For one tandem cell, a thicker (160 nm) LiF antireflective coating (ARC) was evaporated (compared with 100 nm for baseline cells) and led to a stronger perovskite limitation, with photogenerated current densities of 19.20 and 20.76 mA $cm^{-2}$ in the perovskite and silicon, respectively. These cells are particularly suitable for stability measurements because the strong perovskite limitation will immediately affect the performance of the tandem cell if the perovskite subcell changes. The cell was stored in dark in an $N_2$-filled glovebox, and *JV* and MPP (10 min) were regularly measured. After 3000 hours, the cell retained 99.6% of its initial efficiency (29.2%) (fig. S18).

We also reduced the front transparent conductive oxide (TCO) thickness to reduce parasitic absorption and reflection losses, which increased short-circuit current density ($J_{SC}$), and added grid fingers to maintain lateral conductivity and therefore fill factor (FF). We optimized the number of grid fingers and zinc-doped indium oxide (IZO) thickness by performing optical and electrical simulations and then verified the results experimentally. Reducing the IZO from 100 to 40 nm while adding three grid fingers with a pitch of ~2.6 mm enabled a relative PCE gain of ~3%. More details on the optimization can be found in the supplementary materials, materials and methods (figs. S19, S20, and S21 and table S5). A direct comparison of the external quantum efficiency (EQE) and reflection between 100 and 40 nm is shown in Fig. 3B.

These optimizations enable median values (20 cells) of 20.09 mA $cm^{-2}$ $J_{SC}$, 78.27% FF, 1.97 V $V_{OC}$, and a PCE of 31.1% (Fig. 3C and fig. S22). Three-min MPP tracks of all cells are shown in fig. S23 and demonstrate stable devices.

To further improve the optical performance, we decreased the thickness of $C_{60}$ from 18 to 16 nm and $SnO_2$ from 20 to 10 nm. For all tandems presented in this work, we applied a low sputter-damage approach based on a reduced applied radio frequency (RF) power density, which reduces the kinetic energy of damaging particles on the thin $C_{60}$ and $SnO_2$ (*33*). EQE spectra are shown in fig. S24, demonstrating high photogenerated current densities of 20.3 and 21.2 mA $cm^{-2}$ for the perovskite and silicon subcell, respectively. The improvement in $J_{SC}$, although maintaining $V_{OC}$ and FF (Fig. 3C), enabled to increase the PCE to certified 32.5% ± 1.0% (figs. S25 and S26), which corresponds to 76% of the detailed-balance limit. The (quasi) steady-state measurement performed at Joint Research Centre–European Solar Test Installation (JRC-ESTI) is shown in Fig. 3D together with a photo of the tandem device and the performance parameters. The measurement at JRC-ESTI was identical to the in-house measurement (fig. S27).

We performed pseudo-*JV* measurements with subcell-selective intensity-dependent PL and injection-dependent electroluminescence (EL) (figs. S28, S29, and S30 and table S6). QFLS values ($V_{OC}$ values from p*JV*) of ~0.72 and ~1.27 eV were measured for the silicon and perovskite subcell, respectively. The sum of 1.99 eV is in good agreement with the *JV*-measured $V_{OC}$ (1.98 V). The perovskite pseudo-FF of up to 87.8% enables a tandem pseudo-FF of 84.2% and hence a pseudo-PCE of 33.8%. The good agreement between the EL and PL-based p*JV* curves demonstrate well-aligned bands (*12*).

### Discussion

We demonstrated the advantage of PI modification over the commonly implemented LiF interlayer. As revealed with sensitive constant final state yield spectroscopy measurements, the PI did not influence the density of occupied gap states at the interface, whereas LiF induced a higher occupied gap-state density in the $C_{60}$, which is consistent with our previous observations (*22*). This still allows for an improved interface passivation with LiF because this interlayer not only induces a more asymmetric charge carrier concentration at the interface (*22*), but theoretical considerations also suggest that a LiF "spacer" reduces wave function overlap (*17*) and hence the nonradiative across-interface recombination through these additional gap states in the $C_{60}$ (*21*). Under this specific perovskite composition (3Hal) and PI concentration (0.3 mg $ml^{-1}$), the main effect of PI is not a chemical passivation of the perovskite surface but rather allowing for the offset between the conduction band (CB) and lowest unoccupied molecule orbital to be substantially reduced (by ~350 mV). Hence, the quasi–Fermi level of electrons can move closer to the perovskite CB edge. In addition, from the point of view of charge separation, the surface photovoltage measurements show increased electron selectivity at the perovskite surface in the presence of PI. We are aware of the different working mechanism of PI on the perovskite surface compared with that proposed by Li *et al.* (*18*), and we believe that this difference is caused by the different perovskite composition and the concentration of PI used (1 mg $ml^{-1}$ in their work, and 0.3 mg $ml^{-1}$ in this work).

Implementing high-quality perovskite material and using PI as a suitable surface treatment strategy allowed us to achieve an exceptionally high $V_{OC}$ for single junctions (up to 1.28 V, which corresponds to 91.5% of the detailed-balance limit) and tandems (up to 2.0 V). The tandem cells retained 80% of the initial efficiency (28.2%) after $T_{80}$ = 347 hours and 75.7% after 478 hours of continuous MPP tracking. In addition, by combining several optimization processes, we obtained highly efficient tandem devices, with a certified PCE of 32.5%. Subcell-selective measurements revealed a pseudo-PCE of 33.8%. Further improvement could be achieved for a double-side textured cell, for which the improvement in $J_{SC}$ is expected to be ~0.5 mA $cm^{-2}$ (~2.5% relative gain) (*6*).

We provide insight on the working mechanism and consequent improvement of perovskite

and silicon-perovskite solar cell performance using 3Hal with PI. Although our results show a slight improvement of the stability compared with those of reference devices, we believe that further studies may give positive perspectives for future commercialization.

## REFERENCES AND NOTES


1. S. Akhil *et al*., *Mater. Des.* **211**, 110138 (2021).
2. NREL, Best research-cell efficiency chart, Photovoltaic Research; https://www.nrel.gov/pv/cell-efficiency.html.
3. M. H. Futscher, B. Ehrler, *ACS Energy Lett.* **1**, 863–868 (2016).
4. B. Chen *et al*., *Adv. Energy Mater.* **12**, 2003628 (2022).
5. E. Aydin *et al*., *Nat. Energy* **5**, 851–859 (2020).
6. M. Jošt *et al*., *Energy Environ. Sci.* **11**, 3511–3523 (2018).
7. A. Al-Ashouri *et al*., *Science* **370**, 1300–1309 (2020).
8. P. Tockhorn *et al*., *Nat. Nanotechnol.* **17**, 1214–1221 (2022).
9. S. Colella *et al*., *Chem. Mater.* **25**, 4613–4618 (2013).
10. B. Chen *et al*., *Joule* **3**, 177–190 (2019).
11. M. Azhar *et al*., *Mater. Chem. Phys.* **287**, 126335 (2022).
12. F. Lang *et al*., *ACS Energy Lett.* **6**, 3982–3991 (2021).
13. J. Xu *et al*., *Science* **367**, 1097–1104 (2020).
14. C. Pereyra, H. Xie, M. Lira-Cantu, *J. Energy Chem.* **60**, 599–634 (2021).
15. T. Wu, X. Li, Y. Qi, Y. Zhang, L. Han, *ChemSusChem* **14**, 4354–4376 (2021).
16. C. M. Wolff, P. Caprioglio, M. Stolterfoht, D. Neher, *Adv. Mater.* **31**, e1902762 (2019).
17. J. Liu *et al*., *Science* **377**, 302–306 (2022).
18. F. Li *et al*., *J. Am. Chem. Soc.* **142**, 20134–20142 (2020).
19. J. Seo *et al*., *Energy Environ. Sci.* **7**, 2642–2646 (2014).
20. M. Stolterfoht *et al*., *Nat. Energy* **3**, 847–854 (2018).
21. J. Warby *et al*., *Adv. Energy Mater.* **12**, 2103567 (2022).
22. D. Menzel *et al*., *Adv. Energy Mater.* **12**, 2201109 (2022).
23. L. Canil *et al*., *Energy Environ. Sci.* **14**, 1429–1438 (2021).
24. F. Peña-Camargo *et al*., *Appl. Phys. Rev.* **9**, 021409 (2022).
25. T. Dittrich, S. Fengler, *Surface Photovoltage Analysis of Photoactive Materials* (World Scientific, 2020).
26. I. Levine *et al*., *Joule* **5**, 2915–2933 (2021).
27. K. Xu *et al*., *ACS Energy Lett.* **7**, 3600–3611 (2022).
28. M. Khenkin *et al*., *Nat. Energy* **5**, 35–49 (2020).
29. H. Köbler *et al*., *Energy Technol.* **10**, 2200234 (2022).
30. K. Domanski *et al*., *Energy Environ. Sci.* **10**, 604–613 (2017).
31. A. Al-Ashouri *et al*., *ACS Energy Lett.* **8**, 898–900 (2023).
32. A. Cruz *et al*., *Sol. Energy Mater. Sol. Cells* **236**, 111493 (2022).
33. M. Härtel *et al*., *Sol. Energy Mater. Sol. Cells* **252**, 112180 (2023).
34. T. Dittrich, S. Bönisch, P. Zabel, S. Dube, *Rev. Sci. Instrum.* **79**, 113903 (2008).


## ACKNOWLEDGMENTS


The authors thank J. Beckedahl, C. Ferber, M. Wittig, T. Lußky, and H. Heinz for the technical support in HySPRINT Helmholtz Innovation Lab; C. Klimm for SEM technical support; R. MacQueen for optimizing the TRPL setup; and A. Morales Vilches, J. Kleesiek, A. Harter, T. Henschel, and M. Günther for preparing the silicon bottom cells. We gratefully acknowledge T. Dittrich for fruitful discussions and providing the HZB SPV laboratory facilities.

**Funding:** Funding was provided by the German Federal Ministry for Economic Affairs and Climate Action (BMWK). This work was supported by the Helmholtz Association, the HySPRINT Helmholtz Innovation Lab, PVcomB, "Zeitenwende–Zukunftstechnologie Tandem–Solarzellen" as well as the EU Partnering project "TAPAS." We also acknowledge HyPerCells (a joint graduate school of the University of Potsdam and the Helmholtz-Zentrum Berlin) and the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation), project number 423749265–SPP 2196 (SURPRISE) for funding. We acknowledge financial support by the Federal Ministry for Economic Affairs and Energy within the framework of the 7th Energy Research Programme (P3T-HOPE, 03EE1017C). Further funding was provided by the German Federal Ministry for Economic Affairs and Climate Action under grants 03EE1017A (Project P3T), 03EE1086C (PrEsto), and 0324293B (Dynasto). Additional funding was provided by the Federal Ministry of Education and Research through the project "PEROWIN" (Förderkennzeichen: 03SF0631). I.L. thanks the AiF project (ZIM-KK5085302DF0) for financial support. F.L. acknowledges funding by the Volkswagen Foundation through the Freigeist Program. A.Mu. acknowledges financial support from DFG in the framework of the priority program SPP 2196 and funding from the European Union HORIZON-MSCA-2021-PF-01-01 under grant agreement 101061809 (HyPerGreen). M.S. acknowledges the Heisenberg program from DFG for funding—project number 498155101. A.Ma. acknowledges funding from the Research Council of Lithuania under grant agreement 01.2.2-LMT-K-718-03-0040 (SMARTMOLECULES). **Authors contributions:** S.M. conceived the idea, designed, and performed the main experiments and wrote the manuscript. E.K. fabricated tandem devices, introduced the grid fingers, performed stability tests, and assisted in writing the manuscript. F.S. fabricated numerous samples, performed measurements, and supported the data analysis and text writing. K.S. contributed to the conception of the idea, performed essential experiments for the 3Hal optimisation, and provided SEM and EDX data. L.Z. optimized the film thickness of $C_{60}$ and $SnO_2$ and fabricated tandem devices (including record device). M.P. helped performing initials studies on PI used in perovskite single junctions. B.L. helped with the device fabrication strategy and designed the tandem device stability setup (TARIS). F.Y. helped with the optimization of the 3Hal perovskite and performed TRPL analysis. J.W. and F.L. measured PL and pseudo-*JV* analysis of solar cells. A.Mu. measured and analyzed KP and trSPV transients. A.Mu. and I.L. performed analysis of SPV results. D.Men. and S.K. performed XPS and UPS and CFSYS analysis. D.Ma. synthesized the PI molecule. A.A.-A. and A.Ma. provided suggestions for improving the perovskite wettability on Me-4PACz. M.S.H. optimized the IZO deposition process, designed and implemented the low sputter-damage approach and determined *n,k* data required for simulations. K.X. provided the initial idea of using 3Hal perovskite. A.C., P.W. and J.K. provided optimized silicon bottom cells. H.K. performed stability tests on single-junction devices. J.L. measured GIWAXS. D.Mec., E.U., A.A., and M.S. provided funding for this research and gave feedback during the manuscript writing. B.S. provided silicon bottom cell devices fabricated in his laboratory, substantially contributed in optimizing the silicon devices for tandem applications, and provided funding (Dynasto). R.S. supervised the project. L.K. helped with UPS-CFSYS, SPV, and PL analysis and gave deep insight in the data understanding. He also repeatedly helped with the manuscript reviews and provided additional funding (projects P3T and PrEsto). S.A. supervised the project and gave essential suggestions about experimental planning and construction of the manuscript. He also provided the main funding of this project. **Competing interests:** There is no conflict of interest to declare and there are no patents nor patent applications for this work. **Data and materials availability:** All data are available in the main text or the supplementary materials.

**License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS


science.org/doi/10.1126/science.adh5872
Figs. S1 to S30
Tables S1 to S6
References (*35*–*46*)

Submitted 30 November 2022; accepted 4 May 2023
10.1126/science.adf5872

STRUCTURAL BIOLOGY

# Translation dynamics in human cells visualized at high resolution reveal cancer drug action

Huaipeng Xing[1,2], Reiya Taniguchi[1], Iskander Khusainov[1], Jan Philipp Kreysing[1,3], Sonja Welsch[4], Beata Turoňová[1], Martin Beck[1,5]*

Ribosomes catalyze protein synthesis by cycling through various functional states. These states have been extensively characterized in vitro, but their distribution in actively translating human cells remains elusive. We used a cryo–electron tomography-based approach and resolved ribosome structures inside human cells with high resolution. These structures revealed the distribution of functional states of the elongation cycle, a Z transfer RNA binding site, and the dynamics of ribosome expansion segments. Ribosome structures from cells treated with Homoharringtonine, a drug used against chronic myeloid leukemia, revealed how translation dynamics were altered in situ and resolve the small molecules within the active site of the ribosome. Thus, structural dynamics and drug effects can be assessed at high resolution within human cells.

The eukaryotic ribosome (80S) consists of two subunits (60S and 40S) that translate mRNA into proteins (*1*). Purified 80S ribosomes have been extensively studied in vitro, which has elucidated molecular details of translation (*2*, *3*). The translation process can be divided into four main stages: initiation, elongation, termination, and ribosome recycling (*3*). During elongation, the rotation of the 40S, the association of elongation factors, and the translocation of tRNAs are coordinated to synthesize nascent chains (*3*). tRNAs have three canonical binding sites on the ribosome: aminoacyl (A), peptidyl (P), and exit (E) sites (*3*). A noncanonical tRNA binding site called the Z site, located in an extreme position past the E site, has been identified in structures of isolated translationally inactive mammalian ribosomes (*4*). It is speculated that the Z site may represent a late-stage intermediate of tRNA ejection downstream of the E site with similarity to internal ribosome entry site interactions, but the exact physiological relevance remains obscure (*4*).

The 70S ribosome from *Mycoplasma pneumoniae* and 80S ribosome from *Dictyostelium discoideum* and the respective translation elongation processes have been visualized inside cells using cryo–electron tomography (cryo-ET), revealing differences in translation elongation (*5*–*7*). For example, the most abundant elongation intermediate in bacterial *M. pneumoniae* cells was the "A, P" state wherein tRNAs were identified in the respective positions. By contrast, the elongation factor bound "eEF1A, A/T, P" state was most abundant in eukaryotic *D. discoideum* cells. In microsomes isolated from human cells, the membrane-associated elongation cycle has been resolved, revealing the eEF1A, A/T, P, E state as the most prominent (*8*). However, to which extent these insights are applicable to actively translating human cells remains unclear.

Homoharringtonine (HHT) is a nat… compound that binds to the ribosome and … hibits protein synthesis (*9*). It is used to treat patients with chronic myeloid leukemia clinically and as a reference to study new anticancer ribosome inhibitors (*9*, *10*). The HHT-bound ribosome structure has been determined by incubating the purified 80S ribosome with HHT in vitro (*11*, *12*), revealing the binding site at the peptidyl transferase center (PTC). In the cellular context, it remains unclear how HHT affects translation dynamics and how exactly it inhibits protein synthesis. We applied cryo–focused ion beam (cryo-FIB), cryo-ET and advanced data processing algorithms to determine the near-atomic structures of ribosomes and analyzed the abundance and organization of different ribosome states in untreated and HHT-treated human cells.

## Ribosomes are bound with HHT inside human cells

To study the 80S ribosome structures inside human cells, we first prepared cryo–FIB-milled lamellae from 35 native (untreated) human embryonic kidney cells (HEK-293) and acquired 358 tilt series (fig. S1A and table S1). We used template matching to identify ribosomes within the reconstructed tomograms (fig. S1B and

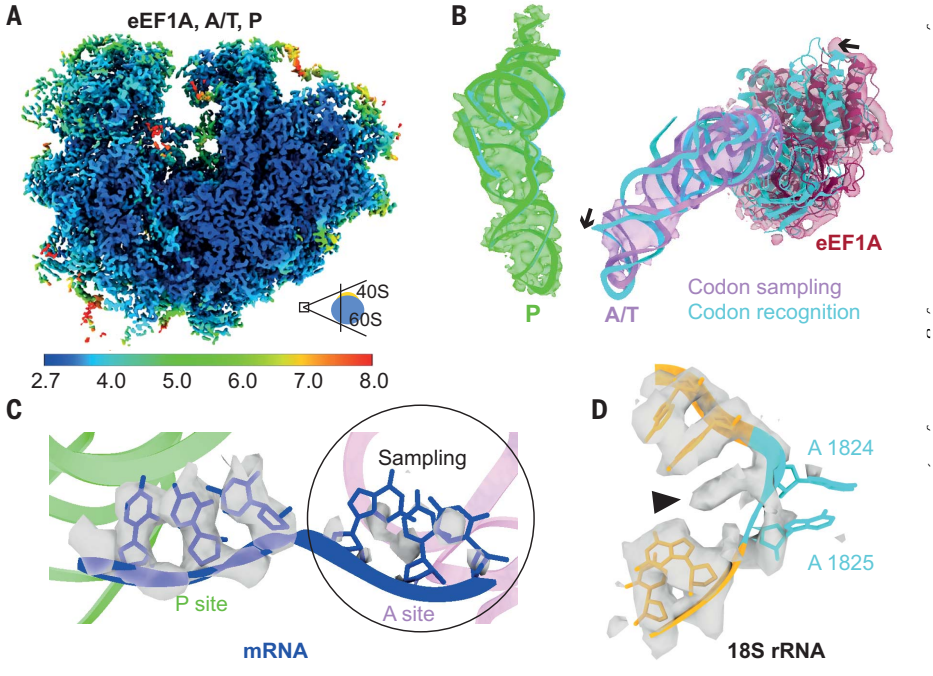



**Fig. 1. 80S ribosome structures in human cells.** (**A**) Color-coded local resolution map of the eEF1A, A/T, P state identified in the untreated dataset. (**B**) Superposition of the codon recognition state (PDB 5LZS, cyan) onto our eEF1A-tRNA ternary complex at the eEF1A, A/T, P state. The atomic model of tRNA (lavender) and eEF1A (maroon) from PDB 4CXG were fitted into our map for comparison to 5LZS. (**C**) The mRNA segmented from the eEF1A, A/T, P state was fitted with PDB 5LZS. Density is more clearly defined at the P site as compared with the A site. (**D**) The atomic model (5LZS) of 18S rRNA (A1822 to A1827) was rigid-body–fitted into the corresponding density of the 80S at the eEF1A, A/T, P state. The decoding nucleotides A1824 and A1825 in the atomic model are in the flipped-out configuration, which is inconsistent with the observed density (arrowhead).

[1]Department of Molecular Sociology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany. [2]Faculty of Biochemistry, Chemistry and Pharmacy, Goethe University Frankfurt am Main, 60438 Frankfurt am Main, Germany. [3]IMPRS on Cellular Biophysics, 60438 Frankfurt am Main, Germany. [4]Central Electron Microscopy Facility, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany. [5]Institute of Biochemistry, Goethe University Frankfurt, Frankfurt am Main, Germany.
*Corresponding author. Email: martin.beck@biophys.mpg.de

movie S1). After classification and refinement, we determined the ribosome structure at an overall resolution of ~3.2 Å and locally resolved features up to 2.5 Å in resolution from untreated cells (Fig. 1B and figs. S1 to S3). The density of ribosomal protein side chains, ribosomal RNA (rRNA) bases, tRNA bases, and ions were well-resolved (fig. S1E), indicating the quality of the map. Using the same approach we processed 352 tilt series from 32 HHT-treated cells. In the resulting structure the density of HHT was visible (fig. S1, F and G) and the position of HHT at the PTC showed a high similarity to previous in vitro analysis (*9*, *11*). The P-tRNA was poorly resolved compared with the untreated ribosome (fig. S1G), suggesting that the translation was altered after HHT treatment.

### Translation elongation cycle in human cells

We next investigated the translation elongation process. In untreated cells, focused classification was performed with dedicated tRNA and elongation factor masks, resulting in eight ribosome states resolved from 3.4 Å to 16.4 Å (figs. S2 and S3 and table S2). These states were well-explained by previously reported structural models that captured various elongation states (figs. S4 to S6) (*4*, *8*, *13*–*15*). The eEF1A, A/T, P state was the best resolved, globally reaching ~3.4 Å (Fig. 1A), and displayed structural features characteristic for tRNA scanning, indicating that codon sampling rather than codon recognition is occurring (Fig. 1B and fig. S4A) (*13*, *16*). Specifically, the mRNA codon was more clearly resolved at the P site (Fig. 1C), suggesting that the codon-anticodon interaction at the A site is not yet established. Consistently, we observe the decoding nucleotides A1824 and A1825, which are thought to stabilize the A-site tRNA by a "flipping out" movement upon codon recognition (*13*), in a "flipped in" conformation (Fig. 1D).

In total, six classes—accounting for ~83% of all identified ribosomes—were assigned to the translation elongation cycle based on their similarity to the previously reported structures (Fig. 2A) (*3*, *5*, *13*, *16*). The elongation cycle is thought to start from the non-rotated P state that was detected but not particularly abundant in our data (*5*) (Fig. 2A). It is followed by the eEF1A, A/T, P state discussed above which was the most prominent, consistent with in situ analysis of the lower eukaryote *Dictyostelium discoideum* (*7*), but in contrast to previous analysis in bacteria (*5*). The abundance of the codon sampling state in human cells could be important for the higher decoding fidelity during translation (*17*). The following states were the non-rotated A/T, P and A, P states. At a low contour level of the map a factor-like density at the eEF1A binding region is seen in the A/T, P state, but neither an extended nor compact eEF1A model provides a sufficient explanation (fig. S4C, Materials and Methods) (*8*, *13*). As expected, the A, P state displayed the A1824 and A1825 nucleotides in a flipped out conformation (fig. S4D). Yet it was less abundant than in bacteria, where it was the most populated (*5*). The subsequent state was identified as eEF2, ap/P, pe/E (*18*). Next, two possible sequences of events are conceivable: eEF2 and E-tRNA could first dissociate from the ribosome forming the P state that subsequently binds the eEF1A-tRNA, or the eEF1A-tRNA may displace eEF2 with subsequent departure of the E-tRNA (*5*) (Fig. 2A). Various conformational changes of the small subunit are apparent during the elongation cycle (fig. S5) and consistent with previous work (*16*, *18*).

Two further states were identified whose places in the elongation cycle are not obvious (Fig. 2B). The rotated eEF2 state without tRNA fitted the previous structural model very well (*15*, *19*) (fig. S4G), although the eEF2 position was shifted by ~1.7 nm compared with the eEF2, ap/P, pe/E state (fig. S4I). The remaining class, which accounts for only 1.6% of all classified ribosomes (Fig. 2B), was consistent with the non-rotated eEF1A, A/T, P, Z state, which has been suggested to be transient (*4*). Although the resolution of this state was only moderate in untreated cells, the observed electron optical density at the Z site fitted well with the purified ribosome at the non-rotated P, Z state (Fig. 2B and fig. S4, H and J) (*4*). To our knowledge, alternative factors binding to this

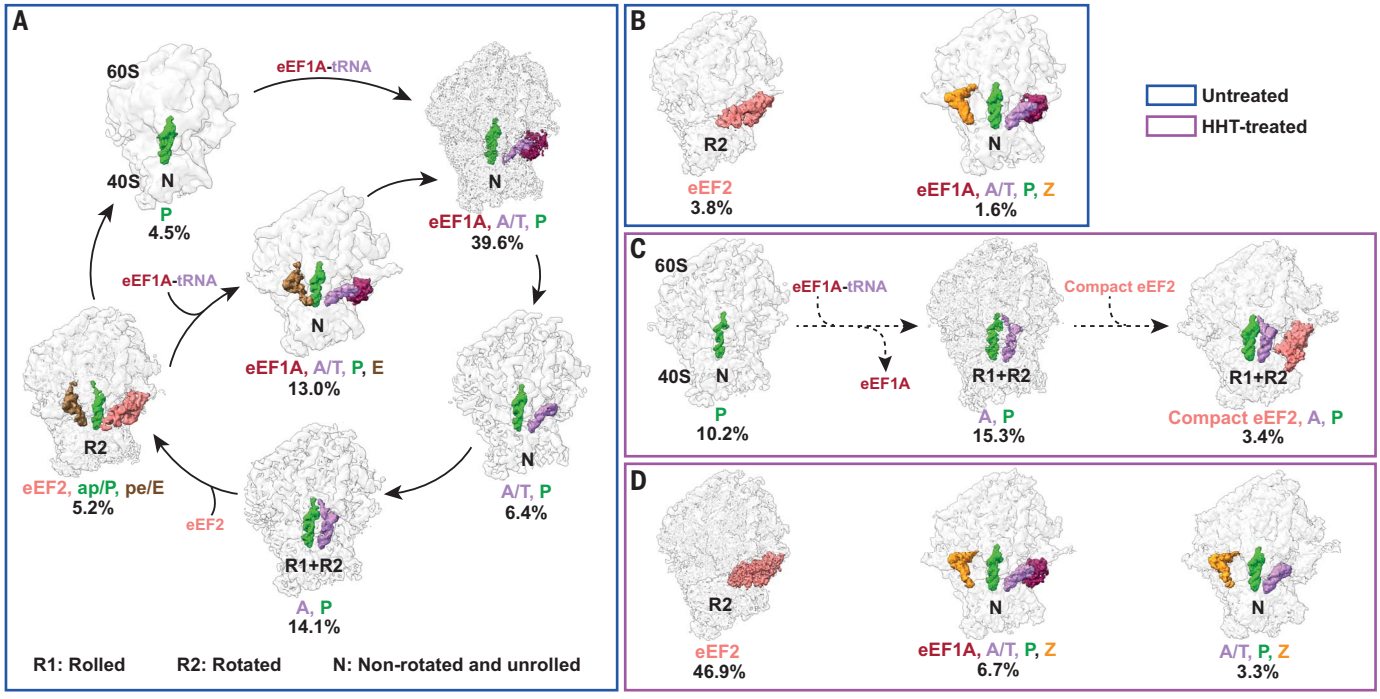



**Fig. 2. Ribosome states in native untreated and HHT-treated cells.** (**A**) Six ribosome states are assigned to the elongation cycle in untreated cells. (**B**) The eEF2 and eEF1A, A/T, P, Z states in untreated cells. (**C**) Three potential translation elongation intermediate states in the HHT-treated cells. The dashed arrow illustrates how the elongation states may connect. (**D**) The eEF2, eEF1A, A/T, P, Z and A/T, P, Z states inside the treated cells. R1, rolled; R2, rotated; N, nonrotated and unrolled. The information about 40S rolling and rotation is shown in figs. S5 and S11.

site have not been identified. More importantly, we obtained higher resolution in HHT-treated cells (see below), which further strengthened this conclusion. We note that ~12% of the particles identified as ribosomes were not attributed to a specific state by our classification and may correspond to additional low abundant states or transitions between states. Collectively, our data allowed the detailed analysis of the translation dynamics inside native human cells, which exhibited differences compared with previous ex vivo studies (*8*, *18*), indicating that the ribosome isolation process may affect the ability to capture native elongation intermediates (*5*, *7*).

### HHT alters the elongation cycle

The anticancer drug HHT blocks peptide bond formation (*9*, *20*). Previous in vitro analysis of purified ribosomes revealed that HHT binds in the A-site cleft in the PTC (*11*, *21*). Upon HHT treatment, one would thus expect to observe a strong enrichment of ribosomes with the A- and P-site tRNAs bound (*11*). Indeed, *M. pneumoniae* ribosomes treated with 0.5 mg/ml (1547.4 μM) chloramphenicol (Cm), a related antibiotic binding to the PTC of the 50S, showed over 70% of ribosomes at the A, P state (*5*, *22*). The effect of HHT on human ribosome states in cells remains unknown. To address this, we exposed cells to HHT concentrations of 0.055 mg/ml (100 μM) in the medium, which resulted in a significantly lower cellular protein concentration than untreated cells (fig. S7A), implying that protein synthesis was inhibited. At the time point examined by cryo-EM, the ATP levels, which serve as an indicator of cell viability, were not yet reduced (fig. S7B) and the cell morphology of untreated and treated cells was indistinguishable (fig. S7C). Classification, as described above, identified six ribosome states resolved into the 3.7 Å to 11.5 Å range (figs. S8 to S12 and table S3). Unexpectedly, the rotated state with eEF2 but without tRNA accounted for almost half of all ribosomes in the dataset (Fig. 2, C and D, and Fig. 3A). This state showed density for SERPINE1 mRNA-binding protein 1 (SERBP1) and thus accounts for hibernating ribosomes (*15*, *23*) (fig. S10G), in which the HHT density was visible (fig. S8C). The P state increased in abundance compared with the untreated cells, whereby the A, P states showed a similar abundance (Fig. 2C). The HHT molecule was also resolved at the A, P state but not in other less populated classes as a result of the lower resolution (Fig. 3B). We therefore combined the particles from the four remaining less abundant classes. The resulting average also displayed density for HHT (fig. S8C). The positions of A-tRNA and P-tRNA at A, P states were similar in untreated and HHT-treated cells (Fig. 3B) (*9*). One of our classes appeared similar to a potential compact eEF2, A, P state (figs. S10, D, I, and J, and S11F) (*24*), although the respective local resolution prevents a definitive assignment due to a lack of secondary structure. Finally, two states with Z-tRNA were considerably more abundant compared with untreated cells (Fig. 2D and fig. S10, C and E), showing the typical features of tRNA shape at the Z site (fig. S10, E and H, and movie S3). Thus, HHT treatment results in the accumulation of ribosome hibernation instead of the A, P state, which may be representative of the mechanism of the drug action in cancer therapy.

To investigate cell-to-cell variability, we analyzed the distribution of the ribosome states across individual cells captured by multiple tomograms. We identified some degree of cell-to-cell variability (fig. S13A), for example, the eEF1A, A/T, P state varied from 25 to 56% in untreated cells (fig. S13B). Overall, the signal observed in multiple tomograms of the same cell was similar, with some notable exceptions that may imply local variability (fig. S13B). Notably, the abundance of the eEF1A, A/T, P (class 1) and eEF1A, A/T, P, E (class 4) was largely anticorrelated in untreated cells (fig. S13, C and D), whereas the sum of eEF1A-bound states was relatively consistent, suggesting that its total concentration may be limited (fig. S13E). By contrast, the abundance of ribosomes containing eEF2 was much more diverse between different cells (fig. S13E).

### Polysomes are impaired upon HHT treatment

We analyzed the spatial organization of different ribosome states within polysomes in treated and untreated cells. The densities accounting for neighboring ribosomes that are commonly observed in the subtomogram averages of untreated cells were reduced in HHT-treated cells (fig. S14A). Furthermore, the ribosome distribution was more dispersed subsequent to treatment (Fig. 4, A and B), both implying that polysomes might be largely abolished. To test this hypothesis, we defined polysomes based on a distance threshold of 9 nm from the mRNA exit to entry sites of neighboring particles (fig. S14B and table S4, Materials and Methods). Although this arbitrary classifier will not capture the entirety of polyribosomes as it could, for example, miss more distantly spaced ribosomes on a longer transcript, overall it performs well in their detection (fig. S15). Of all ribosomes in untreated cells, 30.2% were grouped into arbitrary polysomes (Fig. 4C and fig. S14, C and D), which was considerably higher than in HHT-treated cells (fig. S15, A to E). In monosomes the eEF2 state was prevalent (Fig. 4D and fig. S16, A and B) and exhibits much less neighboring density (fig. S16, C and D), underlining the notion that these ribosomes are hibernating. In polysomes, the eEF2, ap/P, pe/E state (class 3, see Fig. 3A) was less frequent in the first leading ribosome than the trailing ribosomes (fig. S15F). Classification of all untreated ribosomes with a dedicated mask resolved the low-resolution structure of the di-ribosome, in which the mRNA density was visible (fig. S17A) (*25*, *26*). This di-ribosome resembled the top-to-top configuration (t-t, the central protuberance of both ribosomes facing a similar direction) (*27*, *28*) (fig. S17B). Two other arrangements of pairs were apparent: one with the central protuberance of the i+1 ribosome toward down (t-d), the other toward up (t-u) (fig. S17, C and D). Although the abundance of these configurations declined in the treated cells (fig. S17, D and E), the center-to-center and exit-to-entry distance of these pairs were indistinguishable from untreated cells (fig. S17F).

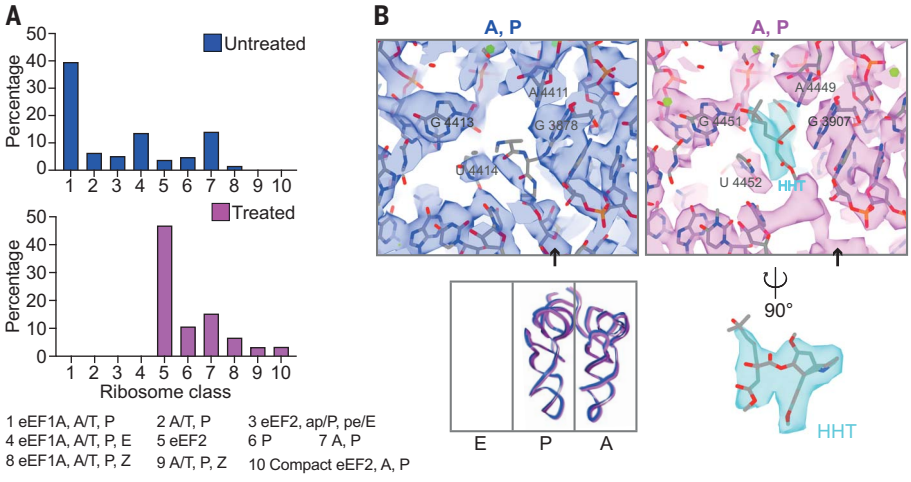



**Fig. 3. Distribution of ribosome states in cells.** (**A**) Percentages of ribosome states in untreated and HHT-treated cells. (**B**) The PTC of ribosomes at the A, P state in untreated and treated cells is shown fitted with PDBs 5AJ0 and 6QZP, respectively. Structural overlay of tRNAs at A, P states with models determined from previous studies (bottom; materials and methods). The black arrow points to P-tRNA. HHT is colored in cyan.

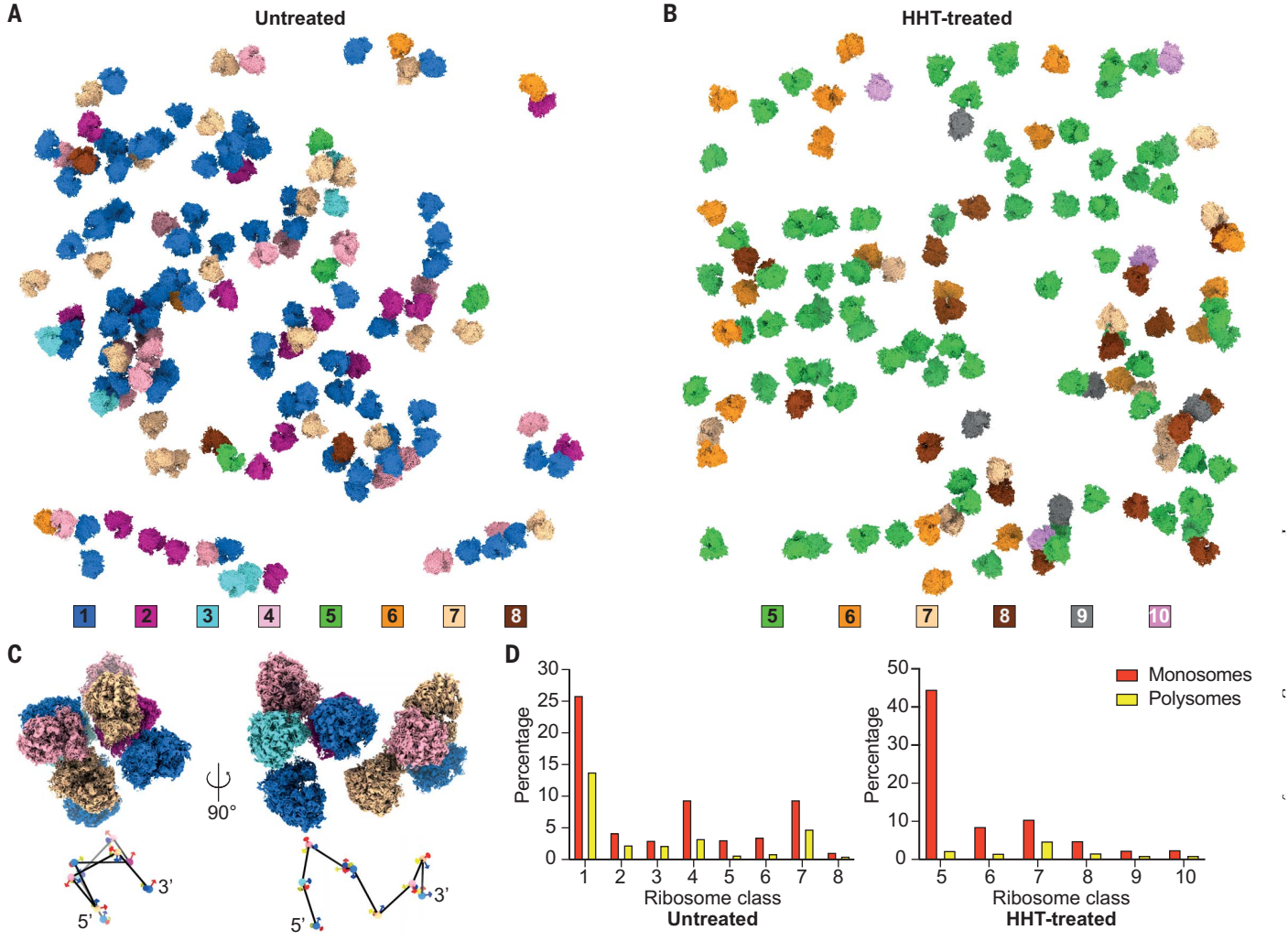



**Fig. 4. Spatial and functional analysis of polysomes.** (**A** and **B**) Distribution of ribosome states in a representative tomogram from an untreated (A) and treated (B) cell. (**C**) A polysome and the putative mRNA path. The center-to-center distance of the neighboring ribosomes: 27.3 ± 1.8 nm (mean ± SD, *n* = 8). (**D**) Distribution of ribosome states in monosomes and polysomes from 358 tomograms of untreated cells and 352 tomograms of treated cells. Class numbers are the same as in Fig. 3A.

We classified ribosomes with a mask covering the region beyond the exit tunnel, thus capturing both the potential membrane and the expansion segment ES27L (figs. S18 and S19, and tables S2 and S3). The analysis revealed that the translation state distributions of soluble and membrane-bound ribosomes were similar (fig. S18D). Notably, seven conformations of the expansion segment ES27L were found in cytosolic ribosomes (fig. S19, Materials and Methods). One of these contains a long stretch of ES27L that associates with the ErbB3 receptor-binding protein (Ebp1) on the surface of the 60S and resembles a previously published in vitro structure (*29*, *30*) (fig. S19, A and B, and tables S2 and S3). The percentage of Ebp1-associated ribosomes was below 20% in most untreated cells, whereas the abundance was over 25% in all HHT-treated cells (fig. S20A). Whether Ebp1 is related to specific translation states remained unclear (*29*, *30*). Thus, a subfraction of Ebp1-bound ribosomes is observed in all apparent intermediates of the elongation cycle (fig. S20B).

### HHT binds the free 60S associated with eIF6

Finally, we investigated the free 60S and 40S in the cytoplasm of human cells. The free 60S showed Ebp1 and ES27L density in both datasets (Fig. 5, A and B, fig. S21, A and B), which was similar to the Ebp1-associated 80S ribosome. However, the 60S was much more abundant in treated cells than in untreated cells (Fig. 5C). Eukaryotic initiation factor 6 (eIF6) binds the 60S to prevent premature 60S binding with 40S (*31*, *32*). Our data show that the 60S did not contain eIF6 in the cytosol of native untreated cells, contrasting HHT-treated cells (Fig. 5D). Thus, it appears that eIF6 prevents the association between large and small ribosomal subunits in cells (*33*–*35*) (Fig. 5, C and D, and fig. S21, C to E). Furthermore, the HHT density was resolved in the 60S from treated cells (Fig. 5E). Thus, we conclude that HHT can bind not only the PTC of different states of the 80S ribosome (Fig. 3B and fig. S8C) but also the free 60S decorated with eIF6, which may block the assembly of 40S and 60S in the cytosol of the cell.

### Discussion

Our extensive cryo-ET analysis stresses the feasibility of obtaining structures at high resolution and visualizes an anticancer drug inside human cells. The local resolution in the ribosome core in our study is limited by the pixel size. This illustrates that technical obstacles that hinder high resolution in cellular tomography have been overcome. The combined improvements in energy filtering (*36*), image processing (Warp-RELION-M pipeline) (*22*) and tight control over specimen thickness (*37*, *38*) enable high resolution, as long as the

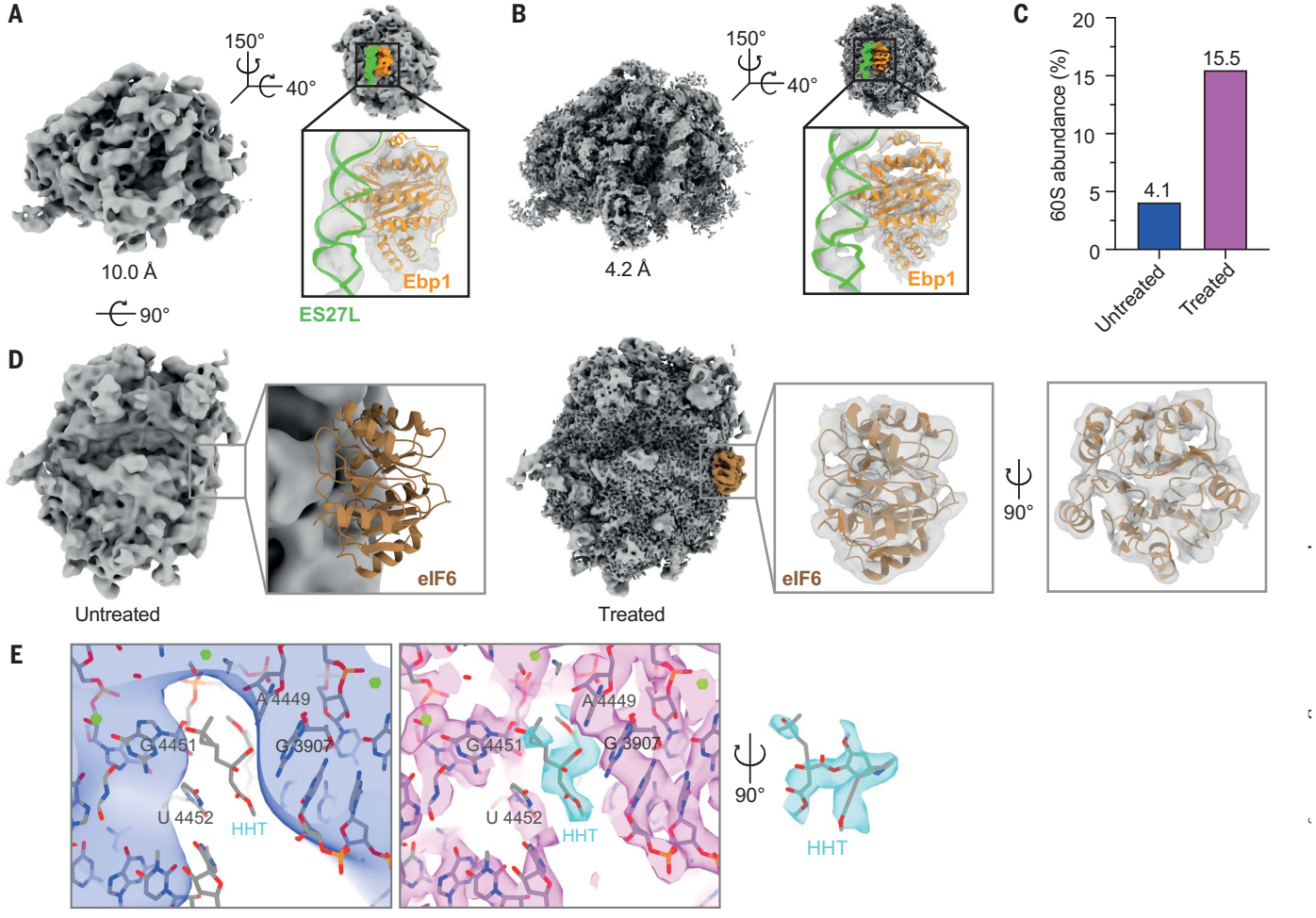



**Fig. 5. Structure of 60S in human cells.** (**A** and **B**) Structures of free 60S in the cytoplasm of untreated cells (A) and HHT-treated cells (B). ES27L and Ebp1 (PDB: 6SXO) are fitted into the 60S maps. Although Ebp1 was not confidently assigned to the 60S from untreated cells as a result of the lower resolution, it fitted better in comparison to alternative factors (amino-terminal acetyltransferases and nascent polypeptide-associated complex) binding to the tunnel exit. (**C**) Percentage of the 60S in untreated and HHT-treated cells normalized to the number of 80S ribosomes in the respective dataset. In untreated cells, abundance = 60S/(60S+80S) = 1,693/(1693+39,402). In treated cells, 60S/(60S+80S) = 7176/(7176+39,070). (**D**) The structure of 60S in untreated cells fitted with PDB 6LSR indicates that eIF6 is missing (left), contrasting HHT-treated 60S (right). Large subunit GTPase 1 (LSG1), NMD3 and ZNF622 (PDB: 6LSR) are not observed in the HHT-treated 60S structure. (**E**) The PTC of free 60S from untreated and treated datasets fitted with PDB 6QZP. HHT is colored in cyan.

number of particles that can be obtained is sufficient. Further improvements in the automatization of specimen preparation techniques may thus be instrumental in pushing the resolution for other macromolecular assemblies analyzed inside cells (*38*–*43*).

The detailed comparison of ribosome structures within untreated and HHT-treated cells revealed that HHT binds to the PTC of the 80S ribosome in situ (Fig. 3B and figs. S1G and S8C), where it can block peptide bond formation as suggested by previous in vitro studies (*9*, *11*, *21*). However, we also observed SERBP1 binding to ribosomes, a factor that functions downstream of mammalian target of rapamycin complex 1 (mTORC1) and leads to a dormant ribosome state (*44*). Such a dormant ribosome state can be a result of cellular stress signaling (*44*). Further, we uncovered that HHT bound the free 60S associated with eIF6 (Fig. 5, D and E), thus likely hampering the assembly of functional 80S ribosomes and in turn further impairing overall protein synthesis. The accumulation of 60S-eIF6 may be explained by the fact that the HHT binding site overlaps with that of the N terminus of Shwachman-Bodian-Diamond syndrome protein (SBDS) (fig. S21F) and that the binding of SBDS to 60S is necessary to release eIF6 (*34*, *45*).

Our study provides insights into the translation elongation cycle, the coordination of ribosome activity within polysomes, and the diverse arrangements of ES27L inside human cells. The same technology may be used in the future to investigate ribosome, translation, and mRNA quality control pathways in the context of human tissue culture cells and primary patient-derived cells. Intermediates that have not yet been observed inside cells may be enriched by introducing kinetical bottlenecks through genetic or pharmaceutical perturbation (*5*, *46*, *47*). Our study may set the stage for the analysis of structures inside mammalian cells and allow characterization of drug susceptibility of human individuals at high resolution by cryo-ET.

## REFERENCES AND NOTES

1. V. Ramakrishnan, *Cell* **159**, 979–984 (2014).
2. A. A. Korostelev, *Annu. Rev. Biochem.* **91**, 245–267 (2022).
3. A. P. Schuller, R. Green, *Nat. Rev. Mol. Cell Biol.* **19**, 526–541 (2018).
4. A. Brown, M. R. Baird, M. C. J. Yip, J. Murray, S. Shao, *eLife* **7**, e40486 (2018).
5. L. Xue *et al.*, *Nature* **610**, 205–211 (2022).
6. F. J. O'Reilly *et al.*, *Science* **369**, 554–557 (2020).
7. P. C. Hoffmann *et al.*, *Nat. Commun.* **13**, 7435 (2022).
8. M. Gemmer *et al.*, *Nature* **614**, 160–167 (2023).
9. A. Gilles *et al.*, *Cells* **9**, 629 (2020).

10. V. Kumar, M. Jyotirmayee, *Mol. Cell. Biochem.* **478**, 1–17 (2023).
11. N. Garreau De Loubresse *et al*., *Nature* **513**, 517–522 (2014).
12. S. K. Natchiar, A. G. Myasnikov, H. Kratzat, I. Hazemann, B. P. Klaholz, *Nature* **551**, 472–477 (2017).
13. S. Shao *et al*., *Cell* **167**, 1229–1240.e15 (2016).
14. R. Buschauer *et al*., *Science* **368**, eaay6912 (2020).
15. J. N. Wells *et al*., *PLOS Biol.* **18**, e3000780 (2020).
16. T. V. Budkevich *et al*., *Cell* **158**, 121–131 (2014).
17. A. Rozov, N. Demeshkina, E. Westhof, M. Yusupov, G. Yusupova, *Trends Biochem. Sci.* **41**, 798–814 (2016).
18. E. Behrmann *et al*., *Cell* **161**, 845–857 (2015).
19. P. R. Smith *et al*., *Nat. Commun.* **12**, 6789 (2021).
20. S. Pellegrino, S. Terrosu, G. Yusupova, M. Yusupov, *Cancers* **13**, 4392 (2021).
21. G. Gürel, G. Blaha, P. B. Moore, T. A. Steitz, *J. Mol. Biol.* **389**, 146–156 (2009).
22. D. Tegunov, L. Xue, C. Dienemann, P. Cramer, J. Mahamid, *Nat. Methods* **18**, 186–193 (2021).
23. P. R. Smith, S. C. Pandit, S. Loerch, Z. T. Campbell, *Trends Biochem. Sci.* **47**, 477–491 (2022).
24. J. Lin, M. G. Gagnon, D. Bulkley, T. A. Steitz, *Cell* **160**, 219–227 (2015).
25. K. Ikeuchi *et al*., *EMBO J.* **38**, e100276 (2019).
26. M. Narita *et al*., *Nat. Commun.* **13**, 6411 (2022).
27. F. Brandt *et al*., *Cell* **136**, 261–271 (2009).
28. F. Brandt, L. A. Carlson, F. U. Hartl, W. Baumeister, K. Grünewald, *Mol. Cell* **39**, 560–569 (2010).
29. K. Wild *et al*., *Nat. Commun.* **11**, 776 (2020).
30. M. L. Kraushar *et al*., *Mol. Cell* **81**, 304–322.e16 (2021).
31. D. W. Russell, L. L. Spremulli, *J. Biol. Chem.* **254**, 8796–8800 (1979).
32. M. Gartmann *et al*., *J. Biol. Chem.* **285**, 14848–14851 (2010).
33. M. Ceci *et al*., *Nature* **426**, 579–584 (2003).
34. F. Weis *et al*., *Nat. Struct. Mol. Biol.* **22**, 914–919 (2015).
35. P. Jaako *et al*., *Nat. Commun.* **13**, 1562 (2022).
36. M. Obr *et al*., *J. Struct. Biol.* **214**, 107852 (2022).
37. A. Rigort, J. M. Plitzko, *Arch. Biochem. Biophys.* **581**, 122–130 (2015).
38. S. Klumpe *et al*., *eLife* **10**, e70506 (2021).
39. M. Smeets *et al*., *Micros. Today* **29**, 20–25 (2021).
40. S. Klumpe *et al*., *Micros. Today* **30**, 42–47 (2022).
41. Z. Wang *et al*., *Science* **375**, eabn1934 (2022).
42. H. Van den Hoek *et al*., *Science* **377**, 543–548 (2022).
43. S. Mosalaganti *et al*., *Science* **376**, eabm9506 (2022).
44. S. Shetty, J. Hofstetter, S. Battaglioni, D. Ritz, M. N. Hall, *EMBO J.* **42**, e112344 (2023).
45. P. Jaako *et al*., *Nat. Commun.* **13**, 1562 (2022).
46. M. Thoms *et al*., *Science* **369**, 1249–1256 (2020).
47. S. Klinge, J. L. Woolford, *Nat. Rev. Mol. Cell Biol.* **20**, 116–131 (2018).
48. turonova/cryoCAT: Initial release, Version v0.1.0, Zenodo (2023); https://zenodo.org/record/7997724.

## ACKNOWLEDGMENTS

We thank P. C. Hoffmann, M. Tuijtel, and T. Majtner from the Department of Molecular Sociology at the Max Planck Institute of Biophysics for advice on membrane segmentation, data collection and data processing. We thank all the members from the Central Electron Microscopy facility and the Department of Molecular Sociology at the Max Planck Institute of Biophysics, and A. Schwarz, E. Schuman and the Department of Synaptic Plasticity at the Max Planck Institute of Brain Research for support and advice. We acknowledge the support from the Max Planck Computing and Data Facility. We thank M. Rodnina, N. Fischer, and S. Boehm for critical reading of the manuscript. **Funding:** M.B. acknowledges funding from the Max Planck Society. **Author contributions:** Conceptualization: H.X. and M.B. Methodology: H.X., B.T., I.K., and J.P.K. Investigation: H.X., R.T., I.K., J.P.K., S.W., and B.T. Visualization: H.X., B.T., and M.B. Funding acquisition: M.B. Project administration: H.X. and M.B. Supervision: H.X. and M.B. Writing – original draft: H.X., B.T., and M.B. Writing – review and editing: H.X., R.T., I.K., J.P.K., S.W., B.T., and M.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The original tilt series have been deposited in the Electron Microscopy Public Image Archive (EMPIAR-11538). EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers: EMD-16721, 16725, 16726, 16727, 16728, 16733, 16734, 16735, 16736, 16737, 16738, 16739, 16740, 16741, 16742, 16743, 16722, 16744, 16748, 16747, 16749, 16750, 16751, 16752, 16754, 16755, 16756, 16757. The codes used for polysome analysis are part of the cryoCAT repository (*48*) and are deposited at Zenodo. All other data are available in the manuscript or supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh1411
Materials and Methods
Figs. S1 to S21
Tables S1 to S4
References (*49*–*64*)
MDAR Reproducibility Checklist
Movies S1 to S3

View/request a protocol for this paper from *Bio-protocol*.

Submitted 3 March 2023; accepted 5 June 2023
10.1126/science.adh1411

ORGANIC CHEMISTRY

# Dearomative ring expansion of thiophenes by bicyclobutane insertion

Huamin Wang[1], Huiling Shao[2]†, Ankita Das[1]†, Subhabrata Dutta[1], Hok Tsun Chan[2], Constantin Daniliuc[1], K. N. Houk[2]*, Frank Glorius[1]*

Skeletal ring enlargement is gaining renewed interest in synthetic chemistry and has recently focused on insertion of one or two atoms. Strategies for heterocyclic expansion through small-ring insertion remain elusive, although they would lead to the efficient formation of bicyclic products. Here, we report a photoinduced dearomative ring enlargement of thiophenes by insertion of bicyclo[1.1.0]butanes to produce eight-membered bicyclic rings under mild conditions. The synthetic value, broad functional-group compatibility, and excellent chemo- and regioselectivity were demonstrated by scope evaluation and product derivatization. Experimental and computational studies point toward a photoredox-induced radical pathway.

Cyclic organic compounds, especially heterocycles, play a key role in pharmaceutical and natural products chemistry, as well as materials science (*1*–*3*). Ring architecture not only influences properties such as lipophilicity, three dimensionality, and scaffold rigidity but also determines molecular function and reactivity (*1*–*3*). Hence, the development of methods to precisely form and manipulate ring systems is of utmost importance in synthetic chemistry. Skeletal ring enlargement can convert common feedstocks into molecularly complex scaffolds and has the potential to induce important changes in the chemical or biological properties of the original compounds (*4*–*6*). Recently, the insertion of one or two atoms into cyclic molecules for skeletal editing, an approach first developed in 1881 (*7*), has received renewed attention (*8*–*16*). Notable contributions from several groups have been reported in the context of skeletal editing by means of boron (*10*), nitrogen (*11*, *12*), or carbon (*13*–*15*) atom insertion (Fig. 1A). Two-step approaches for dearomative ring enlargement of polycyclic arenes through insertion of oxygen (*17*) and carbon (*18*) have also been demonstrated. Moreover, catalytic strategies for two-atom insertions into highly strained cyclic carbonyls have emerged as a useful tool for the synthesis of bridged and fused rings (*19*) (Fig. 1A). Extending this concept to the insertion of small rings into a heterocyclic parent ring ("ring-in-a-heterocycle") would allow the facile construction of bicyclic frameworks (Fig. 1B).

The hetero(bi)cyclic moiety is the basic scaffold of numerous important biologically active natural products (*3*, *20*, *21*). Insertion of C($sp^3$)-rich rings into heterocycles to form bicyclic compounds is appealing because $sp^3$ richness, three dimensionality, and increased conformational rigidity often facilitate the successful design of new drugs (*1*). From a synthetic perspective, bicyclo[1.1.0]butanes (BCBs) are ideal sources of C($sp^3$)-rich rings because of their rigid and defined three-dimensional nature (*22*–*27*). The peripheral cyclization of unsaturated π-bonds with BCBs provides an approach to introduce bicyclic scaffolds in synthetic chemistry (*28*, *29*). Very recently, the cycloaddition of BCBs to strained cyclopropanes, leading to the formation of bicyclo[3.1.1]heptanes, was reported (*30*, *31*). The insertion of BCBs into heterocycles, which exist widely in nature, to forge heterobicyclic compounds has been severely curbed owing to the synthetic challenges such as activation sequence and chemo- and regioselectivity. Considering the applications of thiophenes in medicinal chemistry (*32*, *33*), we envisioned that skeletal ring enlargement of such compounds, by means of radical insertion of BCBs, would provide a privileged platform for the construction of sulfur-based bicyclic scaffolds and be a blueprint for additional efficient ring-in-a-heterocycle enlargement reactions (Fig. 1B).

This abovementioned heterocycle enlargement hypothesis would potentially operate by a cleavage-rebound sequence. A serious synthetic challenge, namely cleaving heterocycles while avoiding the ring opening of the four-membered carbocycles, requires this transformation to be carried out selectively and under mild conditions (Fig. 1B). To address this problem, a radical strategy triggered by photocatalysis can be considered, owing to its mildness and controllable character (*34*–*39*). Because of the electron-rich character and high triplet energy of thiophenes, any activation strategy making use of photoinduced energy transfer was deemed less feasible (*40*–*43*). Given that the excited-state acridinium salt {$[Acr\text{-}Mes_2]^+[BF_4]^-$}* has a high oxidation potential (*44*–*46*), it m be used as a photoredox catalyst for the pr cise activation of the thiophene rather than the BCB to form highly reactive radical cation intermediate **I**, which would attack the BCB and form another intermediate **II** or **III**. Then, eight-membered ring **IV** would be generated through a ring opening–insertion process, followed by the electron-transfer reduction to obtain the target molecule (Fig. 1C). Herein, a photochemical skeletal ring enlargement of thiophenes toward an array of sulfur-based bicyclic rings using BCB as an insertion unit is presented. The synthetic value, broad functional-group compatibility, and outstanding chemo- and regioselectivity of this protocol were demonstrated by scope evaluation, facile millimolar-scale synthesis and synthetic derivatization (Fig. 1D).

## Reaction optimization

We began our investigation with the use of 3-phenylthiophene (**1a**) and bicyclo[1.1.0]butan-1-yl(morpholino)methanone (**2a**) as reaction partners (see supplementary materials for details; fig. S2). After optimization, product **3a** was obtained in 80% yield after reaction under argon atmosphere for 12 hours with acridinium salt as a photocatalyst and acetonitrile (MeCN) as the solvent, and good regioselectivity [**3a:3a'** = 94:6 regioselectivity ratio (r.r.)] was observed (entry 1, fig. S2). We also evaluated the effect of other reaction parameters on this skeletal ring enlargement. Varying the concentration of **1a** resulted in fluctuations in the yields but little effect on regioselectivity (entries 2 to 4, fig. S2). Moreover, changing the concentration of reaction had no effect on the yield (entries 5 and 6, fig. S2). Additionally, other solvents, such as 1,2-dichloroethane (DCE), acetone, tetrahydrofuran (THF), *N*,*N*-dimethylformamide (DMF), and benzotrifluoride ($PhCF_3$) led to a substantial decrease in the reaction yield (entries 9 to 13, fig. S2). We then investigated several photocatalysts and found that only acridinium salt could drive this chemical conversion (entries 14 to 18, fig. S2). Irradiation with visible light and the presence of photocatalyst were essential in this reaction (entries 20 and 21, fig. S2). Under the abovementioned conditions, **3a** was obtained as the predominant regioisomer over **3a'** (fig. S2). By applying a condition-based sensitivity screening approach (*47*) (Fig. 2), we found that the reaction toward compound **3a** is sensitive to high oxygen concentration and water, while being generally tolerant toward higher temperature and varying light intensities (see fig. S6 for details).

## Substrate scope investigation

With the optimal conditions in hand, substrate scope with respect to the BCBs was first evaluated. As shown in Fig. 2, a series of cyclic amide–substituted BCBs proved compatible

[1]Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster (WWU), 48149 Münster, Germany.
[2]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA.
*Corresponding author. Email: glorius@uni-muenster.de (F.G.); houk@chem.ucla.edu (K.N.H.)
†These authors contributed equally to this work.

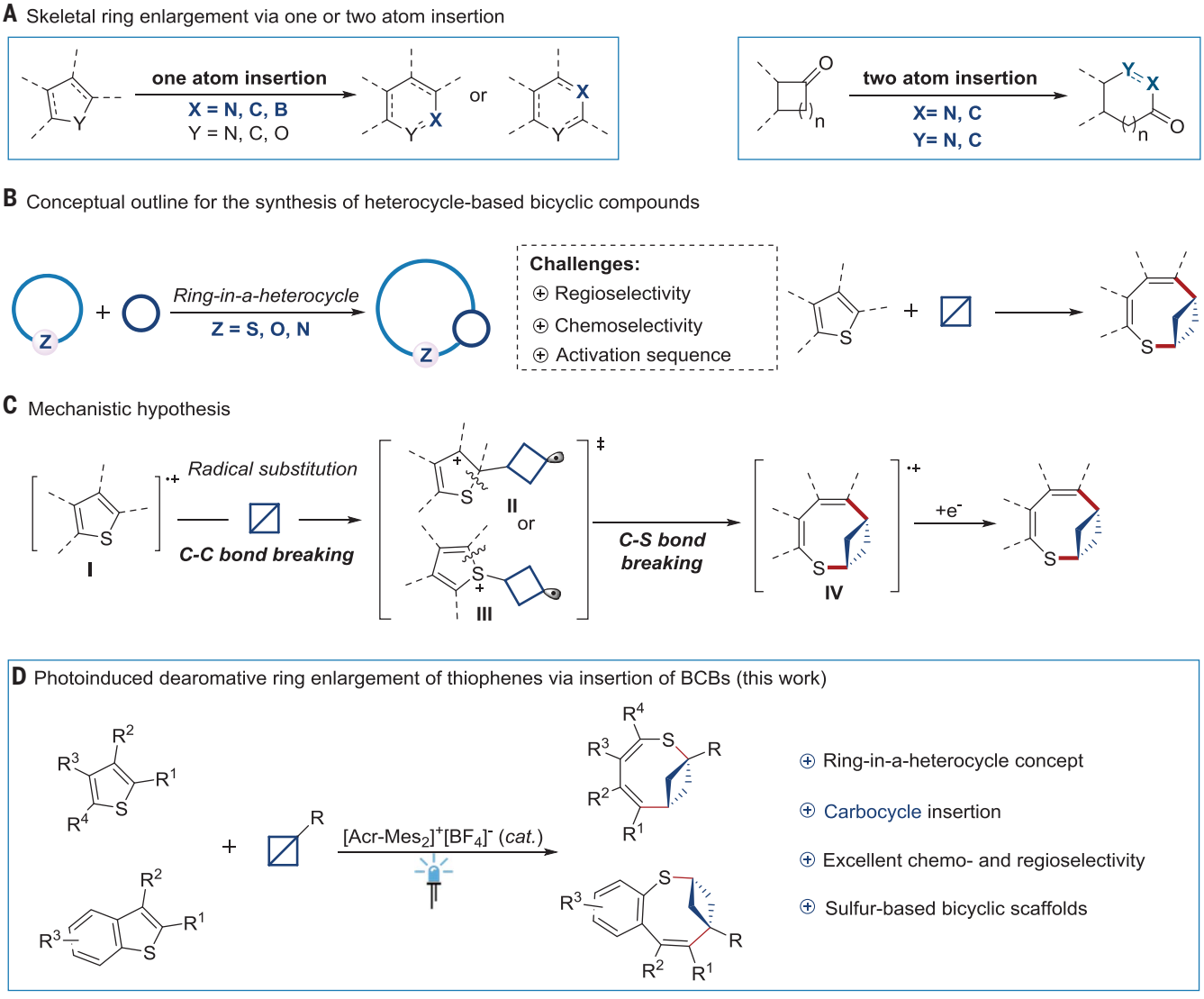



**Fig. 1. Background and concept.** (**A**) Skeletal ring enlargement via one- or two-atom insertion. (**B**) Conceptual outline for the synthesis of heterocycle-based bicyclic compounds. (**C**) Mechanistic hypothesis for the electron transfer–induced insertion of BCBs into thiophenes. (**D**) Photochemical skeletal ring enlargement of thiophenes through insertion of BCBs (this work). cat., catalytic.

with this transformation, affording the corresponding products in good yields and regioselectivities (**3a** to **3e**). The reaction of **1a** with **2a** was successfully performed at millimolar scale, which produced product **3a** in 84% yield. Moreover, other amide groups proved to be compatible, as shown by products **3f** to **3k**. An ester-substituted BCB reacted smoothly with **1a** to afford product **3l**. Additionally, BCBs containing ketone and ether groups gave products in a synthetically useful yield (**3m** and **3n**). Encouraged by the above results, we sought to examine the substrate scope with regard to thiophene derivatives (Fig. 2). High regioselectivities and good yields were observed with monosubstituted thiophenes as reaction partners (**3o** to **3s**). A series of functional groups on the aromatic ring, such as deuterium (**3o**), ketone (**3p**), $CF_3$ (**3q**), fluoride (**3r**), and chloride (**3s**), proved compatible with this protocol, providing many opportunities for further diversifications. These examples show preferential insertion proximal to the aromatic substituent. Further functional-group compatibility of this method was demonstrated with 3,4-disubstituted thiophenes as substrates (**3t** to **3w**). Reactants with aldehyde and silicon-containing ancillary groups were converted to the products **3v** and **3w**. 2,4-Disubstituted thiophenes were subsequently evaluated (**3x** to **3ac**), and higher regioselectivities were obtained than with the 3,4-disubstituted thiophenes, perhaps owing to steric hindrance. Compounds with pyridine substituents were also compatible with the reaction system (**3x** and **3y**). Thiophenes bearing an alcohol or long-chain alkyl group substitution at C2 could be transformed smoothly to the corresponding products through skeletal expansion (**3ab** and **3ac**). This conversion was successfully extended to trisubstituted thiophene, generating compound **3ad** in good yield and superb regioselectivity. Notably, methyl-substituted thiophene was compatible as well (**3ae**). Unsubstituted thiophene reacted in moderate yields to furnish the corresponding products **3af** and **3ag**. Unfortunately, no desired product was detected with the use of 2-methyl-3,5-diphenylthiophene as substrate, which might be due to the steric hindrance caused by the double substitution at the 2,5-position.

Subsequently, the substrate scope of benzothiophene derivatives was explored, and inverted regioselectivities were observed relative

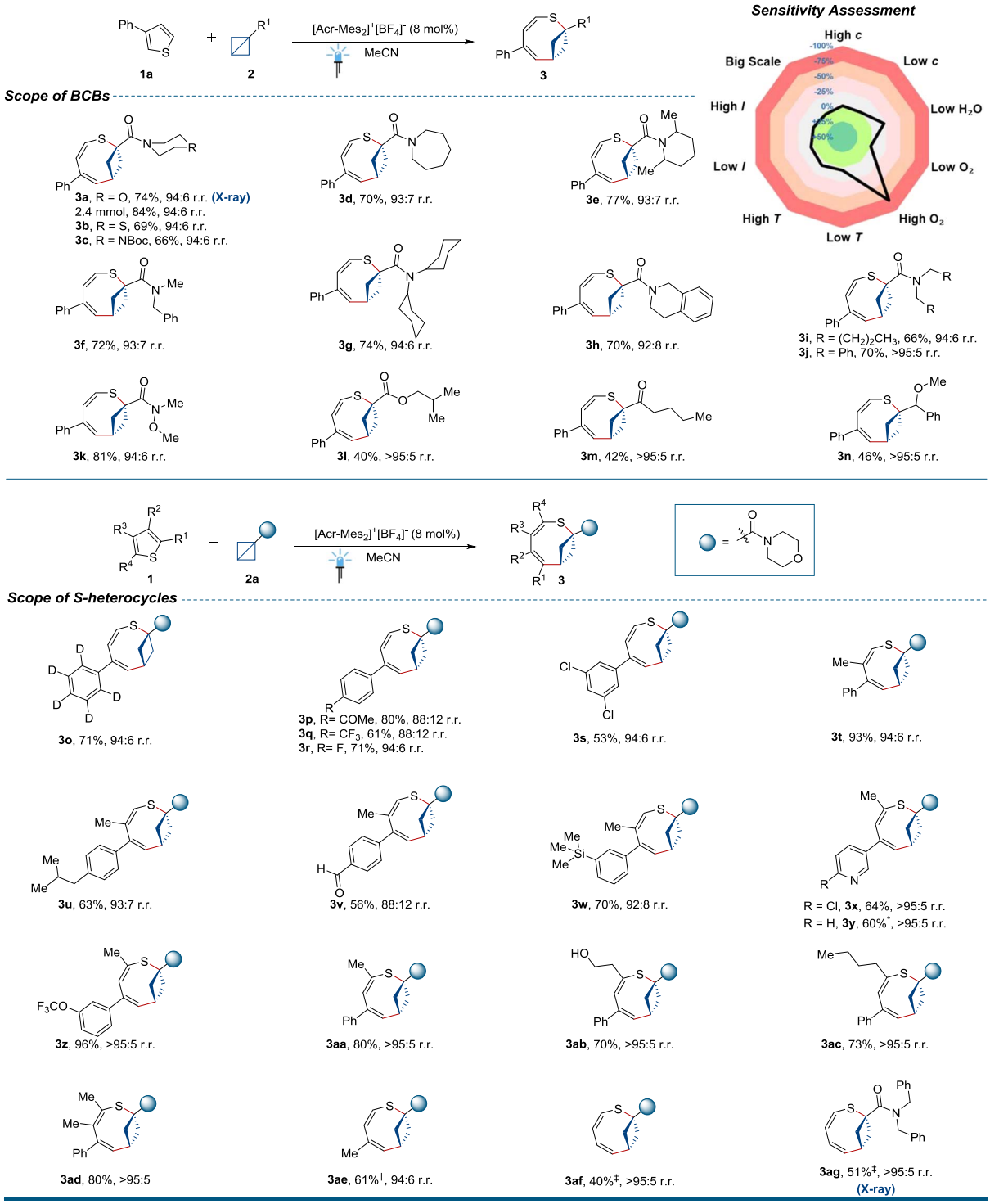



**Fig. 2. Substrate scope for direct insertion of BCBs into thiophene derivatives.** Yields are given for the isolated major products. Regioselectivity ratio is determined by proton nuclear magnetic resonance ($^1$H NMR) analysis of the crude product mixtures. Reaction conditions: thiophene (**1**, 0.36 mmol, 3.0 equiv), BCB (**2**, 0.12 mmol, 1.0 equiv), $[Acr\text{-}Mes_2]^+[BF_4]^-$ (8 mol %), argon, blue light-emitting diodes (LEDs; 450 nm) in MeCN (2 ml) for 12 hours at room temperature; *__1__ (4.0 equiv); †24 hours; ‡$[Acr\text{-}Mes_2]^+[BF_4]^-$ (16 mol %), **1** (4.0 equiv) for 24 hours. *c*, concentration; *T*, temperature; *I*, intensity; Ph, phenyl; Me, methyl; NBoc, *N*-*tert*-butyloxycarbonyl.

to thiophenes (Fig. 3). The target product **3ah** was obtained as a single regioisomer in 59% yield through ring-opening insertion. The robustness of this protocol was demonstrated by its compatibility with various functional groups (e.g., hydroxyl, silicon, phenyl, halogen groups) (**3ai** to **3ao**). A larger-scale reaction successfully afforded product **3ai** in 93% yield. Moreover, the scope was smoothly extended to a thiophene bearing an alkyne (**3ap**). An alkene group, which could have been reactive toward the photosensitizer, also proved compatible under standard conditions, giving product **3av**. Additionally, the cyclohexyl group was well tolerated, as shown by compound **3aq**.

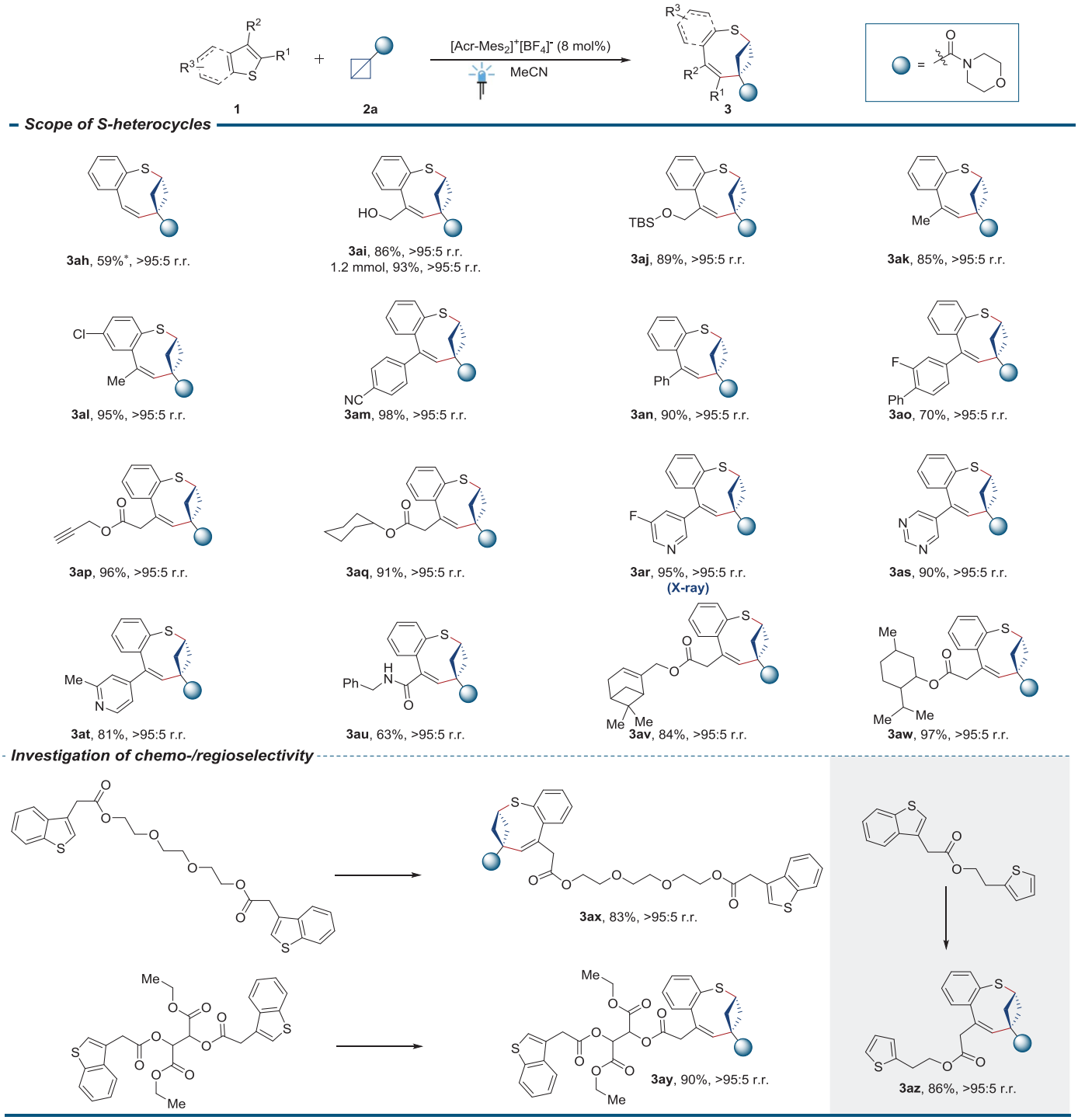



**Fig. 3. Substrate scope for direct insertion of BCBs into benzothiophene derivatives.** Yields are given for the isolated major products. Regioselectivity ratio is determined by $^1$H NMR analysis of the crude product mixtures. Reaction conditions: thiophene (**1**, 0.36 mmol, 3.0 equiv), BCB (**2a**, 0.12 mmol, 1.0 equiv), $[Acr\text{-}Mes_2]^+[BF_4]^-$ (8 mol %), argon, blue LEDs (450 nm), in MeCN (2 ml) for 12 hours at room temperature; *$[Acr\text{-}Mes_2]^+[BF_4]^-$ (16 mol %), **1** (4.0 equiv) for 24 hours. TBS, *tert*-butyldimethylsilyl.

Substrates bearing pyridine and pyrimidine derivatives were amenable to this photochemical protocol, delivering corresponding products in excellent reaction yields (**3ar** to **3at**). Furthermore, benzothiophene bearing an electron-withdrawing group proved to be a suitable partner (**3au**). A substrate containing a menthol fragment was also compatible (**3aw**). To further demonstrate the chemoselectivity and regioselectivity of this heterocycle enlargement, complex molecules with two identical benzothiophene structures were treated with BCB (**2a**). Both reactions led to single-isomer products, respectively, demonstrating exceptional chemo- and regioselectivity (**3ax** and **3ay**). In a substrate with both a benzothiophene and a thiophene moiety, the BCB reacted exclusively at the benzothiophene site (**3az**). Stereochemical and

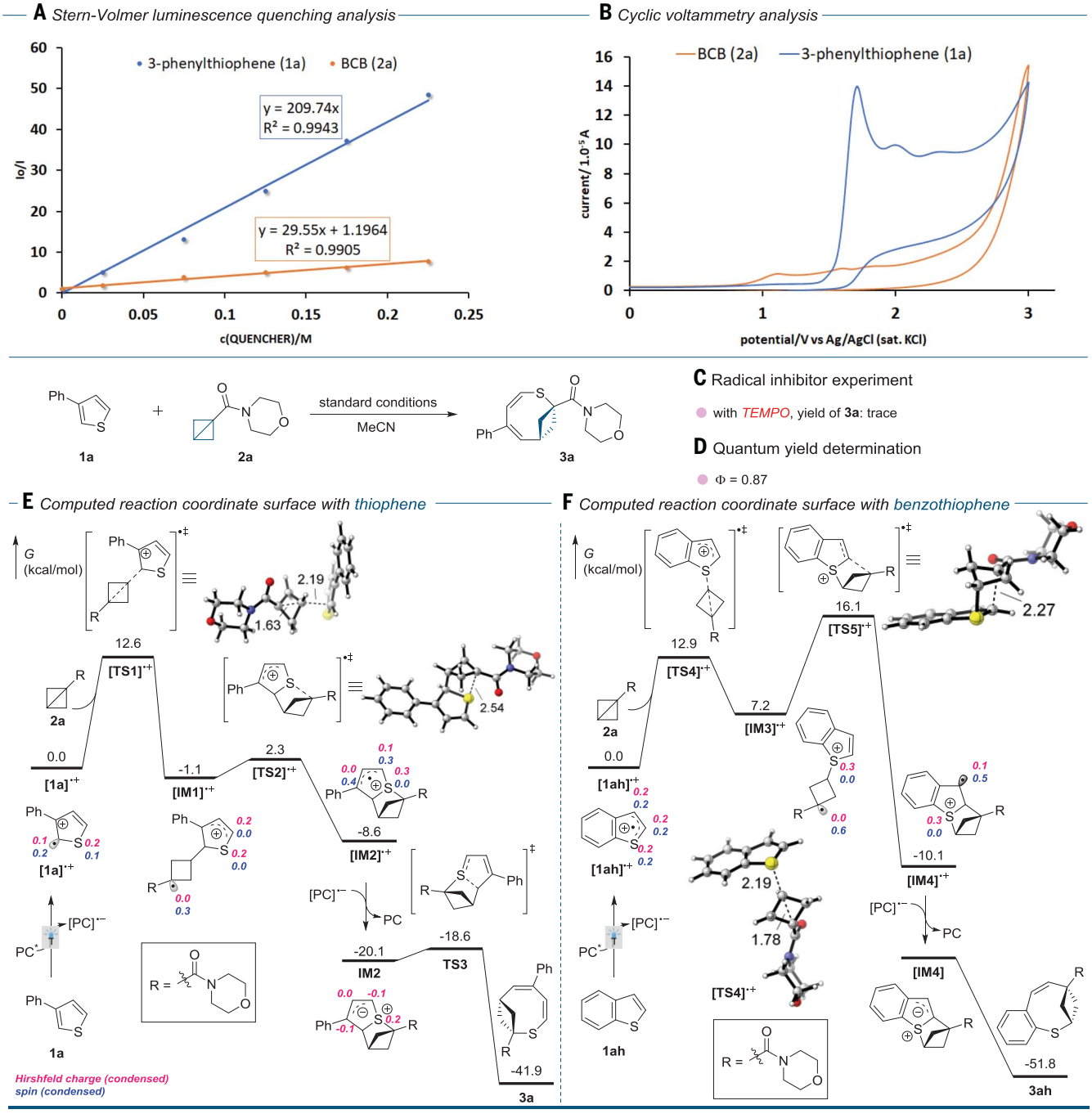



**Fig. 4. Mechanistic studies.** (**A**) Stern-Volmer luminescence quenching analysis. $R^2$, coefficient of determination. (**B**) Cyclic voltammetry analysis. (**C**) Radical inhibitor experiment. (**D**) Quantum yield determination. (**E**) Computed reaction coordinate surface with 3-phenylthiophene. (**F**) Computed reaction coordinate surface with benzothiophene. DFT calculations were conducted at the UM062X/def2TZVP/SMD(solvent = MeCN)//UM06-2X/def2SVP level of theory (*48*–*52*).

regiochemical assignments across the series of products were based on crystal structure analysis of **3a**, **3ag**, and **3ar** (CCDC 2250075-2250077).

### Mechanistic investigation

We then turned our attention to the mechanism of this photoinduced ring enlargement. Stern-Volmer luminescence quenching analysis revealed that the excited-state photocatalyst (PC*) is more easily quenched by thiophene **1a** than by BCB **2a**, consistent with our original mechanistic hypothesis (Fig. 4A). Cyclic voltammetry was then performed to study the redox potentials of **1a** and **2a** (Fig. 4B). An oxidation peak of **1a** in MeCN was observed at 1.68 V [versus saturated calomel electrode (SCE)], and no oxidation peak of **2a** was detected. These results are again consistent with oxidation of **1a** but not **2a** by the PC* (oxidation potential $E_{ox}$ = 2.00 V versus SCE) (*46*). Moreover, no desired product was generated in the presence of the radical scavenger 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO), thereby suggesting a radical pathway (Fig. 4C). The quantum yield Φ of the reaction between **1a** with **2a** was determined to be 0.87 (Fig. 4D). Further intermolecular competition

experiments indicated that electron-rich heterocycles are more reactive than electron-poor heterocycles, selectively providing corresponding products (see supplementary materials for details).

Additionally, we performed density functional theory (DFT) calculations to gain deeper insight into the ring enlargement mechanism (Fig. 4, E and F, and figs. S9 and S10). The experimental studies suggest that the PC* oxidizes **1a** to form the radical cation intermediate [**1a**]$^{\bullet+}$. Our DFT calculation revealed the highest spin density on C2 and the highest Hirshfeld charge on the benzylic carbon, which rationalize the observed BCB insertion at C2 (Fig. 4E). The BCB insertion [**TS1**]$^{\bullet+}$ is the rate-limiting transition state with a free-energy barrier of 12.6 kcal/mol and offers the thermodynamically more stable [**IM1**]$^{\bullet+}$. Subsequent C–S bond formation to produce the heterocycle ([**TS2**]$^{\bullet+}$) is both kinetically facile and thermodynamically favored. Reduction of [**IM2**]$^{\bullet+}$ by the photocatalyst is thermodynamically favored by 11.5 kcal/mol. The final C–S bond cleavage (**TS3**) is kinetically facile, and the generation of product ***3a*** is irreversible.

Alternatively, because benzothiophene expansion takes place with inverted regioselectivity to forge ***3ah***, we hypothesized that the spin density distribution of the radical cation intermediate [**1ah**]$^{\bullet+}$ is different from that of [**1a**]$^{\bullet+}$. Indeed, our calculation showed that the spin density and Hirshfeld charges are evenly distributed over the *S*-heterocycle. Moreover, the reaction mechanism with benzothiophene (Fig. 4F) is different than that with 3-phenylthiophene (**1a**). The BCB insertion (**TS4**) into the *S*-radical has a similar free-energy barrier of 12.9 kcal/mol, as compared with **TS1**. The formation of **INT-5** is endergonic by 7.2 kcal/mol. Furthermore, the subsequent C–C bond formation is the rate-limiting transition state with a free-energy barrier of 16.1 kcal/mol. Generation of the fused heterocycle **INT-6** radical cation is exergonic by 10.1 kcal/mol. Although we could not locate the reduced [**IM4**], we expect the C–S bond cleavage to obtain the product ***3ah*** to be kinetically facile.

## Product elaboration

Next, synthetic derivatization of the eight-membered sulfur-based bicyclic motifs was investigated (Fig. 5; see supplementary materials for details). Starting from compound ***3a***, aldehyde ***3ba*** was successfully produced in good yield through selective amide reduction with THF as solvent. Since sulfone and sulfoxide groups are widely used in organic synthesis, chemoselective oxidation of ***3a*** was carried out and successfully gave sulfoxide ***3bb*** and sulfone ***3bc***. Tertiary amine ***3bd*** could be produced in 64% yield through the reduction of the amide carbonyl group with the preservation of the carbon-carbon double bond. Notably, selective hydrogenation of alkene and diene within ***3a*** offered desired products ***3be*** and ***3bf*** in 76% and 87% yields, respectively, which provides more sulfur-based bicyclic compounds with different properties. Subsequently, with the use of compound ***3ai*** as a starting material, substitution with 3-bromoprop-1-ene provided allyl ether ***3bg*** by introducing an important allyl group. Compound ***3ai*** was also converted to aldehyde ***3bh*** and complex ester ***3bi*** by means of selective oxidation and esterification, respectively. These successful transformations of various functional groups further illustrate the potential applicability of these sulfur-based bicyclic motifs.

Given their wide applications in synthetic and medicinal chemistry, we used thiophene derivatives as the reaction substrates to start this ring-opening insertion protocol. Additionally, because medium-sized bicyclic scaffolds with a nitrogen and oxygen are also prevalent in both naturally occurring and synthetic bioactive molecules, we believe that this ring-in-a-heterocycle protocol will facilitate and inspire

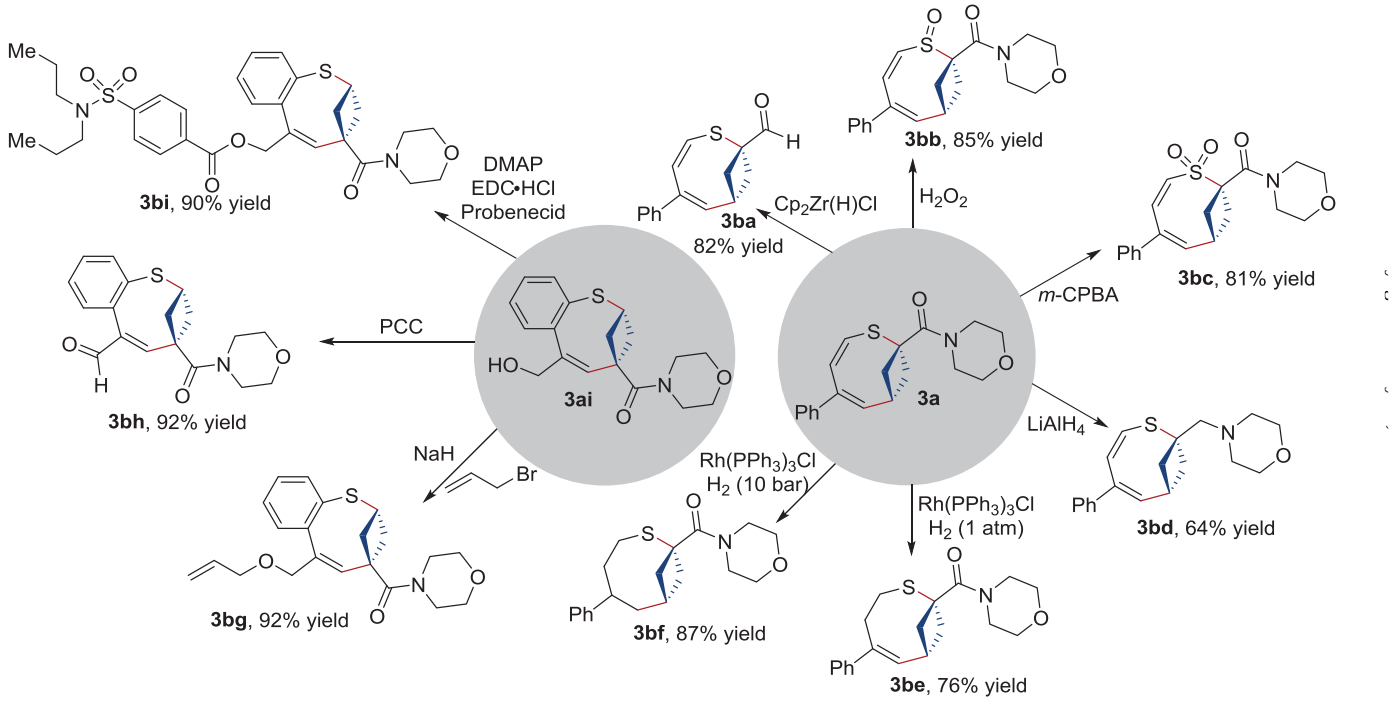



**Fig. 5. Synthetic derivatization.** Reaction conditions for **3ba**: $Cp_2Zr(H)Cl$ (0.2 mmol), **3a** (0.1 mmol), THF (2.5 ml); for **3bb**: $H_2O_2$ (0.125 mmol), **3a** (0.1 mmol), dichloromethane (DCM, 0.2 ml), trifluoroacetic acid (TFA, 0.2 ml); for **3bc**: 3-chloroperoxybenzoic acid (*m*-CPBA, 0.4 mmol), **3a** (0.1 mmol), DCM (2 ml); for **3bd**: $LiAlH_4$ (0.744 mmol), **3a** (0.1 mmol), THF (2 ml); for **3be**: $Rh(PPh_3)_3Cl$ (0.015 mmol), **3a** (0.1 mmol), $H_2$ (1 atm), PhMe (0.5 ml), 9 hours; for **3bf**: $Rh(PPh_3)_3Cl$ (0.015 mmol), **3a** (0.1 mmol), $H_2$ (10 bar), PhMe (0.5 ml), 5 hours; for **3bg**: NaH (0.096 mmol), allyl bromide (0.12 mmol), **3ai** (0.08 mmol), DMF (1.3 ml); for **3bh**: pyridinium chlorochromate (PCC, 0.15 mmol), **3ai** (0.1 mmol), DCM (0.5 ml); and for **3bi**: **3ai** (0.1 mmol), probenecid (0.12 mmol), 4-dimethylaminopyridine (DMAP, 0.12 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC·HCl, 0.15 mmol), DCM (0.5 ml).

the development of skeletal ring enlargements of such heterocyclic compounds through insertion of small rings.

## REFERENCES AND NOTES


1. R. D. Taylor, M. MacCoss, A. D. G. Lawson, *J. Med. Chem.* **57**, 5845–5859 (2014).
2. P. D. Leeson, B. Springthorpe, *Nat. Rev. Drug Discov.* **6**, 881–890 (2007).
3. Y. Chen, C. Rosenkranz, S. Hirte, J. Kirchmair, *Nat. Prod. Rep.* **39**, 1544–1556 (2022).
4. B. Biletskyi *et al*., *Chem. Soc. Rev.* **50**, 7513–7538 (2021).
5. J. R. Donald, W. P. Unsworth, *Chem. Eur. J.* **23**, 8780–8799 (2017).
6. P. Dowd, W. Zhang, *Chem. Rev.* **93**, 2091–2115 (1993).
7. G. L. Ciamician, M. Dennstedt, *Ber. Dtsch. Chem. Ges.* **14**, 1153–1163 (1881).
8. J. Jurczyk *et al*., *Nat. Synth.* **1**, 352–364 (2022).
9. B. W. Joynson, L. T. Ball, *Helv. Chim. Acta* **106**, e202200182 (2023).
10. H. Lyu, I. Kevlishvili, X. Yu, P. Liu, G. Dong, *Science* **372**, 175–182 (2021).
11. J. C. Reisenbauer, O. Green, A. Franchino, P. Finkelstein, B. Morandi, *Science* **377**, 1104–1109 (2022).
12. J. Wang, H. Lu, Y. He, C. Jing, H. Wei, *J. Am. Chem. Soc.* **144**, 22433–22439 (2022).
13. E. E. Hyland, P. Q. Kelly, A. M. McKillop, B. D. Dherange, M. D. Levin, *J. Am. Chem. Soc.* **144**, 19258–19264 (2022).
14. Z. Wang, L. Jiang, P. Sarró, M. G. Suero, *J. Am. Chem. Soc.* **141**, 15509–15514 (2019).
15. M. Mortén, M. Hennum, T. Bonge-Hansen, *Beilstein J. Org. Chem.* **11**, 1944–1949 (2015).
16. S. Tsuchiya, H. Saito, K. Nogi, H. Yorimitsu, *Org. Lett.* **19**, 5557–5560 (2017).
17. Z. Siddiqi, W. C. Wertjes, D. Sarlah, *J. Am. Chem. Soc.* **142**, 10125–10131 (2020).
18. P. Piacentini, T. W. Bingham, D. Sarlah, *Angew. Chem. Int. Ed.* **61**, e202208014 (2022).
19. Y. Xue, G. Dong, *Acc. Chem. Res.* **55**, 2341–2354 (2022).
20. A. Trabocchi *et al*., *J. Med. Chem.* **53**, 2502–2509 (2010).
21. M. Ruiz, P. López-Alvarado, G. Giorgi, J. C. Menéndez, *Chem. Soc. Rev.* **40**, 3445–3454 (2011).
22. K. B. Wiberg, S. T. Waddell, *J. Am. Chem. Soc.* **112**, 2194–2216 (1990).
23. K. Chen, X. Huang, S. B. J. Kan, R. K. Zhang, F. H. Arnold, *Science* **360**, 71–75 (2018).
24. R. Gianatassio *et al*., *Science* **351**, 241–246 (2016).
25. X. Zhang *et al*., *Nature* **580**, 220–226 (2020).
26. W. Dong *et al*., *Nat. Chem.* **14**, 1068–1077 (2022).
27. A. Fawcett, T. Biberger, V. K. Aggarwal, *Nat. Chem.* **11**, 117–122 (2019).
28. R. Kleinmans *et al*., *Nature* **605**, 477–482 (2022).
29. S. Agasti *et al*., *Nat. Chem.* **15**, 535–541 (2023).
30. Y. Zheng *et al*., *J. Am. Chem. Soc.* **144**, 23685–23690 (2022).
31. T. Yu *et al*., *J. Am. Chem. Soc.* **145**, 4304–4310 (2023).
32. R. Shah, P. K. Verma, *Chem. Cent. J.* **12**, 137–158 (2018).
33. E. A. Ilardi, E. Vitaku, J. T. Njardarson, *J. Med. Chem.* **57**, 2832–2842 (2014).
34. C. K. Prier, D. A. Rankic, D. W. C. MacMillan, *Chem. Rev.* **113**, 5322–5363 (2013).
35. K. L. Skubi, T. R. Blum, T. P. Yoon, *Chem. Rev.* **116**, 10035–10074 (2016).
36. F. Juliá, T. Constantin, D. Leonori, *Chem. Rev.* **122**, 2292–2352 (2022).
37. N. Holmberg-Douglas, D. A. Nicewicz, *Chem. Rev.* **122**, 1925–2016 (2022).
38. M. D. Kärkäs, J. A. Porco Jr., C. R. J. Stephenson, *Chem. Rev.* **116**, 9683–9747 (2016).
39. L. Zhang, E. Meggers, *Acc. Chem. Res.* **50**, 320–330 (2017).
40. Q.-Q. Zhou, Y.-Q. Zou, L.-Q. Lu, W.-J. Xiao, *Angew. Chem. Int. Ed.* **58**, 1586–1604 (2019).
41. J. Großkopf, T. Kratz, T. Rigotti, T. Bach, *Chem. Rev.* **122**, 1626–1653 (2022).
42. F. Strieth-Kalthoff, M. J. James, M. Teders, L. Pitzer, F. Glorius, *Chem. Soc. Rev.* **47**, 7190–7202 (2018).
43. W. M. Flicker, O. A. Mosher, A. Kuppermann, *Chem. Phys. Lett.* **38**, 489–492 (1976).
44. I. A. MacKenzie *et al*., *Nature* **580**, 76–80 (2020).
45. C. Song *et al*., *Angew. Chem. Int. Ed.* **58**, 12206–12210 (2019).
46. L. Pitzer, F. Sandfort, F. Strieth-Kalthoff, F. Glorius, *Angew. Chem. Int. Ed.* **57**, 16219–16223 (2018).
47. L. Pitzer, F. Schäfers, F. Glorius, *Angew. Chem. Int. Ed.* **58**, 8572–8576 (2019).
48. Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215–241 (2008).
49. F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
50. A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **113**, 6378–6396 (2009).
51. G. Luchini, J. V. Alegre-Requena, I. Funes-Ardoiz, R. S. Paton, *F1000 Res.* **9**, 291–304 (2020).
52. M. J. Frisch *et al*., Gaussian 16 Revision C.01 (Gaussian Inc., 2016).


## ACKNOWLEDGMENTS

We thank Y. Li (Wuhan University) and R. Mayska, P. Bellotti, and J. E. Erchinger (WWU) for assistance and discussions. **Funding:** Generous financial support was provided by the Alexander von Humboldt Foundation (H.W.), the Deutsche Forschungsgemeinschaft (Leibniz Award), and the International Graduate School for Battery Chemistry, Characterization, Analysis, Recycling and Application (BACCARA). This work was also funded by the Ministry for Culture and Science of North Rhine Westphalia, Germany (A.D.). DFT calculations were performed on the IDRE Hoffman2 cluster at the University of California, Los Angeles, and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation (OCI1053575). Additional funding was provided to K.N.H. by the NSF (CHE-2153972). **Author contributions:** F.G. and H.W. conceived of the project. H.W., A.D., and S.D. designed and conducted experiments. K.N.H., H.S., and H.T.C. conducted the computational studies. C.D. analyzed the x-ray structures. H.W., H.S., K.N.H., and F.G. supervised the research and wrote the manuscript, with contributions from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The Cambridge Crystallographic Data Centre (CCDC) website contains the supplementary crystallographic data for this paper, which can be obtained free of charge at www.ccdc.cam.ac.uk/ data_request/cif under identifiers 2250075, 2250076, and 2250077. All other data are available in the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh9737
Materials and Methods
Supplementary Text
Figs. S1 to S10
X-Ray Data
NMR Spectra
References (*53*–*73*)

Submitted 26 March 2023; accepted 30 May 2023
10.1126/science.adh9737

QUANTUM SIMULATION

# Direct observation of nonlocal fermion pairing in an attractive Fermi-Hubbard gas

**Thomas Hartke*, Botond Oreg, Carter Turnbaugh, Ningyuan Jia, Martin Zwierlein**

**The Hubbard model of attractively interacting fermions provides a paradigmatic setting for fermion pairing. It features a crossover between Bose-Einstein condensation of tightly bound pairs and Bardeen-Cooper-Schrieffer superfluidity of long-range Cooper pairs, and a "pseudo-gap" region where pairs form above the superfluid critical temperature. We directly observe the nonlocal nature of fermion pairing in a Hubbard lattice gas, using spin- and density-resolved imaging of ~1000 fermionic potassium-40 atoms under a bilayer microscope. Complete fermion pairing is revealed by the vanishing of global spin fluctuations with increasing attraction. In the strongly correlated regime, the fermion pair size is found to be on the order of the average interparticle spacing. Our study informs theories of pseudo-gap behavior in strongly correlated fermion systems.**

Long-range Cooper pairs form in a Fermi gas for even the weakest attraction between fermions. With increasing interaction, fermion pairs become more tightly bound, as the system undergoes a smooth crossover from Bardeen-Cooper-Schrieffer (BCS) superfluidity toward a Bose-Einstein condensate (BEC) of molecular pairs (*1*–*3*). In the BCS limit, pair formation and the onset of superfluidity occur at the same temperature, but in the crossover, pairs are expected to form at temperatures above the critical temperature $T_c$ for superfluidity; the onset pair-formation temperature is usually called $T^*$. In this so-called "pseudo-gap" regime, the pair size should be on the order of the interparticle spacing and pairing strongly affected by many-body effects (*4*, *5*). The character of this strongly correlated regime, situated between a Fermi liquid and a normal Bose liquid, is a matter of debate, whose resolution should affect understanding of other strongly coupled fermion systems, such as the high-$T_c$ cuprates and twisted bilayer graphene (*6*–*8*). The rich physics of the BEC-BCS crossover is captured by the attractive Fermi-Hubbard model, a spin-1/2 gas of fermions hopping on a lattice with on-site interactions between unlike spins (*9*–*17*). Through a particle-hole transformation, it stands in one-to-one correspondence with the repulsive Hubbard model (*18*, *19*), which under charge doping is believed to hold the key toward understanding high-temperature superconductivity. The model can be realized using neutral fermionic atoms in optical lattices with tunable interactions. Recent investigations have found spectral gaps (*20*), correlations between local pairs (*21*), and evidence for interspin correlations from density profiles (*22*).

In this work, we observe the formation and spatial ordering of nonlocal fermion pairs in the pseudo-gap regime of an attractive Hubbard gas confined to two dimensions. We use bilayer quantum gas microscopy to detect the in situ location and spin of each fermion in every experimental shot (*23*–*25*). Access to microscopic spin and density correlations reveals the formation of nonlocal pairs, the development of long-range spatial correlations between pairs, and the interplay of pair fluctuations with this density-wave order.

Department of Physics, MIT–Harvard Center for Ultracold Atoms, and Research Laboratory of Electronics, MIT, Cambridge, MA 02139, USA
***Corresponding author. Email: hartke@mit.edu**

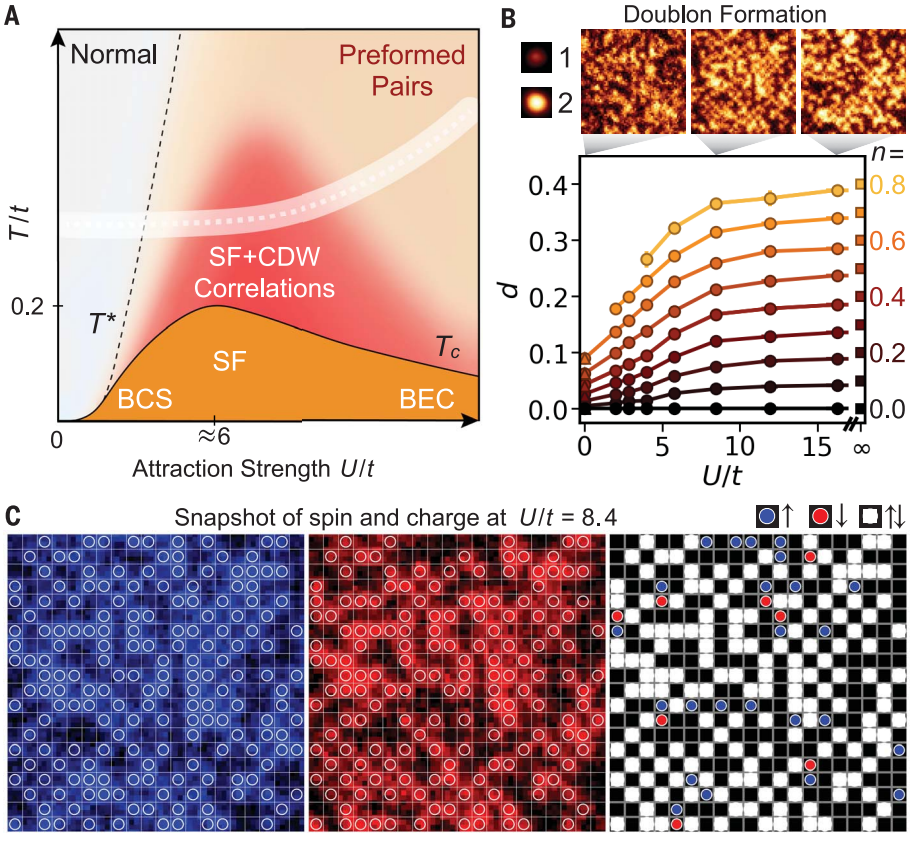



**Fig. 1. Atom-resolved detection of an attractive Fermi-Hubbard gas.** (**A**) Qualitative phase diagram of the attractive Fermi-Hubbard model versus on-site attraction $U/t$ and temperature $T/t$ at density $n \approx 0.8$ (*9*–*13*). Below a critical temperature $T_c$, attractive fermions form a BCS or BEC superfluid (SF). In the pseudo-gap regime between $T_c$ and pairing temperature $T^*$, accessed in this work (white shading), increasing attraction drives pair formation, with pairs exhibiting charge density wave (CDW) and superfluid correlations. (**B**) Measured doublon density $d$ (circles) at fixed density $n$ versus $U/t$, from the noninteracting limit $d = (n/2)^2$ (triangles) to the fully paired limit $d = n/2$ (squares), with representative images of the full density in ~20 × 20 site regions shown above. (**C**) Snapshot of full spin-and-density readout of a strongly correlated gas at $U/t = 8.4(4)$ and $T/t = 0.36(5)$. The spin up (blue), spin down (red), and combined images (rightmost panel) are obtained via bilayer quantum gas microscopy (*23*, *25*).

### Simulating the attractive Fermi-Hubbard model

The phase diagram of the attractive Fermi-Hubbard model is shown in Fig. 1A as a function of the attractive on-site interaction strength $U$, tunneling amplitude $t$, and temperature $T$ (*9*–*13*). For weak attraction $U \ll t$, a BCS superfluid of long-range fermion pairs forms with $T_c = T^*$, reflecting the exponentially weak pair binding. In the opposite limit of strong attraction $U \gg t$, all fermions are bound into local on-site pairs below a dissociation temperature $T^* \sim U$. These pairs condense at the critical temperature of Bose-Einstein condensation $T_c$, proportional to the pair density $n_p$ and pair tunneling rate $t_p \sim t^2/U$. A peak of the condensation temperature $T_c/t \approx 0.2$ is expected to occur at $U/t \approx 6$ and density $n \approx 0.8$ (*9*–*13*). Above the transition temperature, superfluid correlations compete with the formation of a checkerboard charge density wave (CDW) (*21*). At half-filling (density $n = 1$), this competition persists down to $T = 0$ and prevents condensation. In this work we use a filling $n \approx 0.8$, staying in a regime where the ground state is a paired superfluid (*9*, *15*).

We here experimentally realize the attractive Fermi-Hubbard model using a two-species gas of degenerate fermionic $^{40}$K atoms trapped within an optical lattice (*23*, *25*). The Hubbard tunneling amplitude $t$ is controlled by the lattice depth, and the interaction strength $U$ is tuned via the magnetic field. A bilayer quantum gas microscope reveals the location of each fermion.

As a first measure of strong pairing in the attractive Hubbard gas, we detect the density $d$ of doublons (doubly occupied lattice sites) for increasing interaction strength $U/t$ across the phase diagram in Fig. 1A. At fixed density $n$, $d$ increases from the noninteracting limit $d = (n/2)^2$ of random encounters of unlike spins to the fully paired limit $d = n/2$ (Fig. 1B) (*17*). At intermediate attraction, strong checkerboard ordering of doublons is observed, shown in Fig. 1C at $U/t = 8.4(4)$ and $T/t = 0.36(5)$.

Multiple neighboring sites containing a single spin up and spin down are present among doublons in Fig. 1C. These correlated pairs of single spins are evidence of the nonlocal nature of fermion pairs. The microscopic mechanism is the virtual dissociation of a doublon into spatially separate pairing partners, with matrix element $t$ and intermediate energy cost $U$, which perturbatively lowers the energy of a pair by $4t^2/U$. Because pairs are composed of fermions, dissociation can only occur if a nearby site does not already contain a like spin. This leads to effective nearest-neighbor repulsive interactions between pairs (*23*), which in turn are the source of long-range CDW order. The presence of these delocalized pairs also demonstrates that the doublon density $d$ is an incomplete measure of pairing.

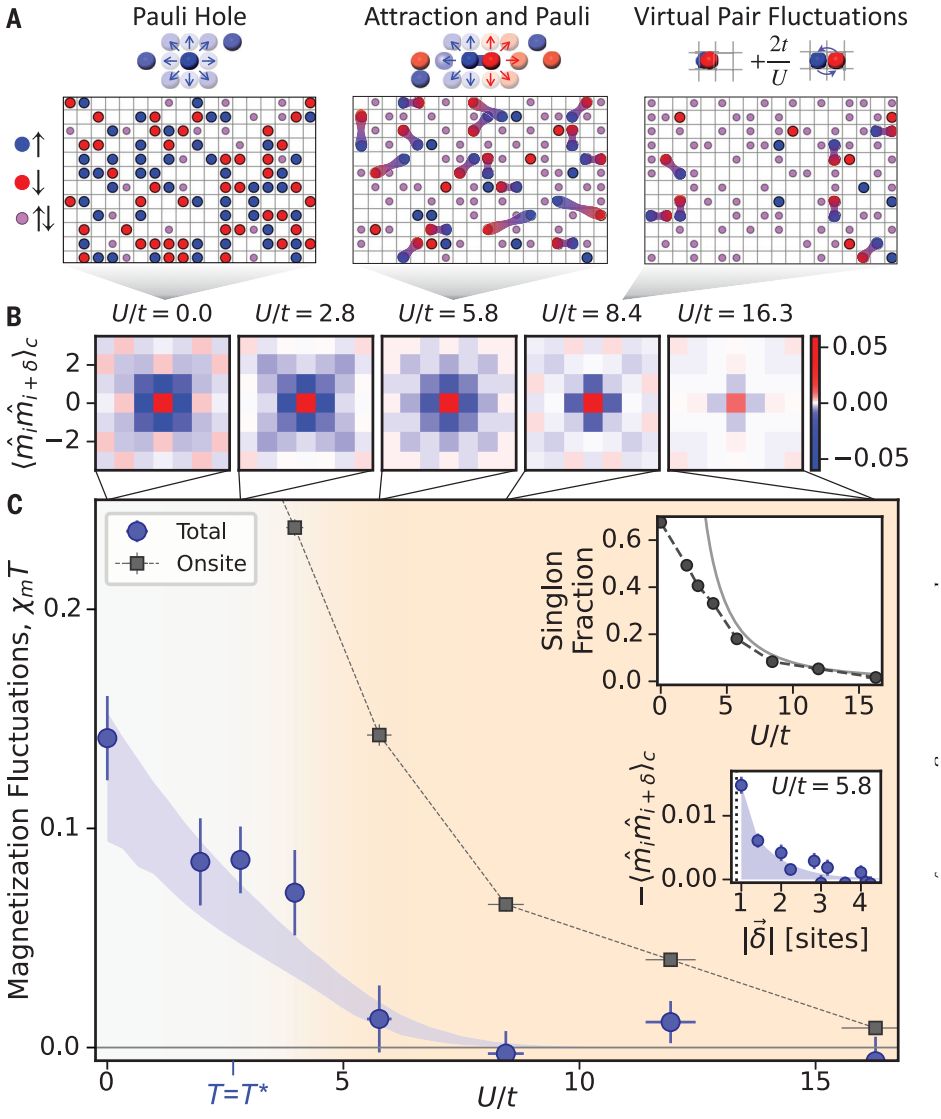



**Fig. 2. Observation of nonlocal fermion pairing.** (**A**) Experimental snapshots of the Fermi gas at $U/t = 0$, $U/t = 5.8(3)$, and $U/t = 8.4(4)$ (left to right), showing the formation of nonlocal pairs and on-site pairs with increasing attraction. Schematics above highlight the physics dominating spin correlations in each image, and shaded bonds suggest possible pair correlations. (**B**) Correlation maps $\langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$ of the magnetization $m = n_\uparrow - n_\downarrow$ at various $U/t$. (**C**) Total magnetization fluctuations $\sum_{\vec{\delta}} \langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$ (blue circles) and on-site fluctuations (black squares) versus $U/t$. Total fluctuations equal the product of magnetic susceptibility $\chi_m$ and temperature $T$ via the fluctuation-dissipation theorem (*23*). Vanishing total spin fluctuations for $U/t \geq 6$ (orange shading) indicate full pairing and vanishing $\chi_m$. Blue shading shows quantum Monte Carlo simulations of total fluctuations at $n = 0.85$, from $T/t = 0.3$ to $T/t = 0.4$ (*23*). The pairing temperature $T^*$ crosses $T \approx 0.35t$ at $U/t \approx 2.5$. Nonlocal pairing is reflected in the singlon fraction per total density $s/n$ (upper inset), which scales as $\sim 8t^2/U^2$ (gray line) at large attraction. Fluctuations at $U/t = 5.8(3)$ (lower inset) extend beyond the interparticle spacing $1/\sqrt{\pi n_\uparrow}$ (dotted line). All data and error bars are obtained from bootstrapping greater than 50 images of atomic clouds with imaging loss correction (*23*).

### Detecting fermion pairing

A true signature of pairing that accounts for these nonlocal pairs is the vanishing of total spin fluctuations. Indeed, a system in contact with a surrounding particle bath will generally display fluctuations of the total magnetization $M = \sum_i \langle \hat{m}_i \rangle$, where the magnetization $m = n_\uparrow - n_\downarrow$. However, pair formation suppresses spin fluctuations, as pairs do not contribute to $M$, and thus in a fully paired system the variance $\sigma_M^2$ vanishes. This variance

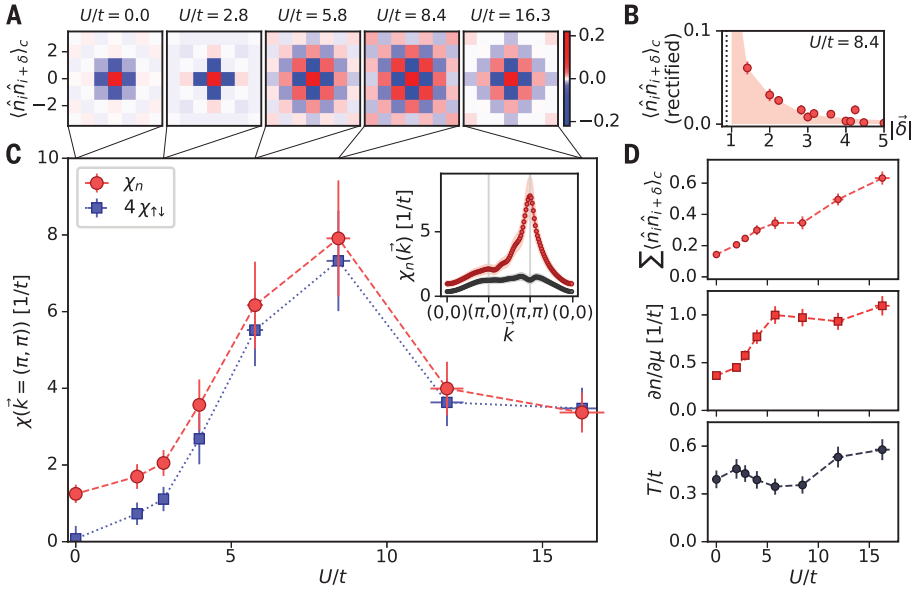



**Fig. 3. Charge density wave ordering of pairs.** (**A**) Density correlations $\langle \hat{n}_i \hat{n}_{i+\delta} \rangle_c$ reflect the crossover from a noninteracting gas (left), to a fully paired gas with CDW order (center), to a weakly ordered gas of local pairs (right). (**B**) At $U/t = 8.4(4)$, the nonlocal rectified correlations $\langle \hat{n}_i \hat{n}_{i+\delta} \rangle_c (-1)^{\delta_x+\delta_y}$ are well described by long-range exponential decay. Dotted line shows the interparticle spacing $1/\sqrt{\pi n_\uparrow}$. (**C**) The density response $\chi_n$ (red circles) and interspin response $\chi_{\uparrow\downarrow}$ (blue squares) at wave vector $\vec{k} = (\pi, \pi)$ serve as order parameters for CDW correlations. These susceptibilities are obtained via the fluctuation-dissipation theorem $\chi_n\left(\vec{k}\right) = (1/T) \sum_{\vec{\delta}} \langle \hat{n}_i \hat{n}_{i+\delta} \rangle_c \cos\left(\vec{k} \cdot \vec{\delta}\right)$ and $\chi_{\uparrow\downarrow}\left(\vec{k}\right) = (1/T) \sum_{\vec{\delta}} \langle \hat{n}_{i\uparrow} \hat{n}_{i+\delta\downarrow} \rangle_c \cos\left(\vec{k} \cdot \vec{\delta}\right)$ (*23*). Inset: $\chi_n\left(\vec{k}\right)$ versus $\vec{k}$ at $U/t = 0$ (black) and $U/t = 8.4(4)$ (red). (**D**) Fluctuation thermometry: The measured density fluctuations $\sum_{\vec{\delta}} \langle \hat{n}_i \hat{n}_{i+\delta} \rangle_c$ (top) and uniform compressibility $\partial n/\partial \mu = \chi_n\left(\vec{k} = (0,0)\right)$ (middle) are combined to directly obtain the temperature $T/t$ (bottom).

is measured locally in our quantum gas microscope through the sum of connected correlations $\sigma_M^2/(\text{Area}) = \sum_{\vec{\delta}} \langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$, where $\langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c = \langle \hat{m}_i \hat{m}_{i+\delta} \rangle - \langle \hat{m}_i \rangle \langle \hat{m}_{i+\delta} \rangle$.

The magnetization fluctuations are directly connected to the magnetic susceptibility $\chi_m = \partial m/\partial h$, the response of the magnetization to a global magnetic field $h$, through the fluctuation-dissipation theorem $\chi_m T = \sum_{\vec{\delta}} \langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$ (*23*). An energy gap for spin excitations, which exponentially suppresses excess spins and thus $\sum_{\vec{\delta}} \langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$, also exponentially suppresses $\chi_m$ (*26*).

Figure 2 reports a crossover to full fermion pairing beyond an interaction strength $U/t \approx 6$ at $n = 0.8(1)$ and $T/t = 0.35(5)$, determined by in situ observation of magnetization fluctuations. The reduction in fluctuations is in good agreement with theoretical predictions for these parameters (*12*–*14*). Figure 2A highlights the physical mechanisms that determine spin fluctuations at various $U/t$. At vanishing interactions, Pauli exclusion separately reduces the density fluctuations of each spin, and thereby also reduces total spin fluctuations. With increasing attraction, nonlocal pairs form in which spins are subject to a competition of Pauli exclusion and attraction, while deep in the on-site pair regime spin fluctuations reflect virtual hopping onto neighboring sites. From statistical averages over more than 50 spin configurations, as shown in Fig. 2A for each interaction strength, we obtain the two-dimensional (2D) magnetization correlation maps $\langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$, shown in Fig. 2B. To detect pairing, Fig. 2C presents the sum of these correlation maps, the total magnetization fluctuations, which are fully suppressed beyond $U/t \approx 6$. Already at zero interactions, Pauli exclusion reduces total fluctuations by $68(5)\%$ compared to the high-temperature expectation $n(1 - n/2)$. This reflects the substantial degeneracy of the Fermi gas ($T/T_F \approx 0.1$, where $T_F$ is the Fermi temperature). Increasing attraction reduces magnetization fluctuations further, and the fraction of unpaired spins is less than $1.5(1.8)\%$ at $U/t = 5.8(3)$, where $\sum_{\vec{\delta}} \langle \hat{m}_i \hat{m}_{i+\delta} \rangle_c$ gives the density of unpaired spins. This full suppression is dual to the formation of a Mott insulator for repulsive interactions (*18*, *19*, *23*, *25*).

The suppression of fluctuations in Fig. 2C with increasing $U/t$ signifies the development of an energy gap for spin excitations (*26*). Theory predicts (*10*–*14*, *27*, *28*) a pairing temperature $T^* \approx 0.25U$ in the crossover regime (*23*). This predicted $T^*$ crosses $T \approx 0.35t$ near $U/t \approx 2.5$, explaining the near-complete suppression of fluctuations beyond $U/t \approx 6$. The corresponding expected spin excitation gap far exceeds the two-body binding energy $E_b$, highlighting the many-body nature of pairing.

Within this regime of full pairing, the two atoms within a pair have a finite probability of being located on separate lattice sites because of quantum fluctuations, in which case they appear as isolated single atoms (called singlons). This probability is measured by the ratio of the singlon density to the total density, $s/n$ (Fig. 2C, upper inset). The observed scaling of $s/n$ with $t^2/U^2$ at strong attraction is expected from perturbation theory already for a Fermi-Hubbard double well (*25*, *29*). At $U/t = 5.8(3)$, the nonlocal portion of the pairs amounts to ~20%. The effective size of fermion pairs can be obtained as the spatial extent of nonlocal spin fluctuations. With full pairing at $U/t = 5.8(3)$, spin fluctuations are present beyond the single-spin interparticle spacing (Fig. 2C, lower inset), indicating that fermion pairs overlap substantially.

### Measuring charge correlations

Characteristic for the pseudo-gap regime is a predicted strong departure from Fermi liquid behavior, in which spin and charge fluctuations are similar (*4*, *5*). Having established the existence of nonlocal fermion pairs through vanishing magnetization fluctuation, we therefore now explore charge (i.e., density) correlations of the gas. Whereas for weak interactions charge and spin correlations are closely associated, for stronger attraction we instead find spatial ordering into a CDW across the phase diagram of Fig. 1A. Previously, evidence for CDW order has only been observed in doublon-doublon correlations and at a fixed interaction strength (*21*). In Fig. 3A, beginning without interactions, we observe negative nonlocal density correlations $\langle \hat{n}_i \hat{n}_{i+\delta} \rangle_c$ for nearest-neighbor and diagonal correlations. These correlations are a simple and direct manifestation of the Pauli hole experienced by a single spin (*25*, *30*). Indeed, here $\langle \hat{n}_i \hat{n}_{i+\delta} \rangle_c$ is precisely equal to $\langle \hat{n}_{i\uparrow} \hat{n}_{j\uparrow} \rangle_c + \langle \hat{n}_{i\downarrow} \hat{n}_{j\downarrow} \rangle_c$ because the measured interspin correlations $\langle \hat{n}_{i\uparrow} \hat{n}_{j\downarrow} \rangle_c$ vanish (*23*). For increasing attraction, the Pauli hole gives way to the positive checkerboard long-range density correlations, shown versus distance in Fig. 3B at $U/t = 8.4(4)$. Further increase in $U/t$ reduces the observed CDW strength, likely as a result of smaller effective repulsion between more-localized pairs (Fig. 2B). Notably, we measure a negative sum of nonlocal interspin correlations $\langle \hat{n}_{i\uparrow} \hat{n}_{i+\delta\downarrow} \rangle_c$ for any attractive

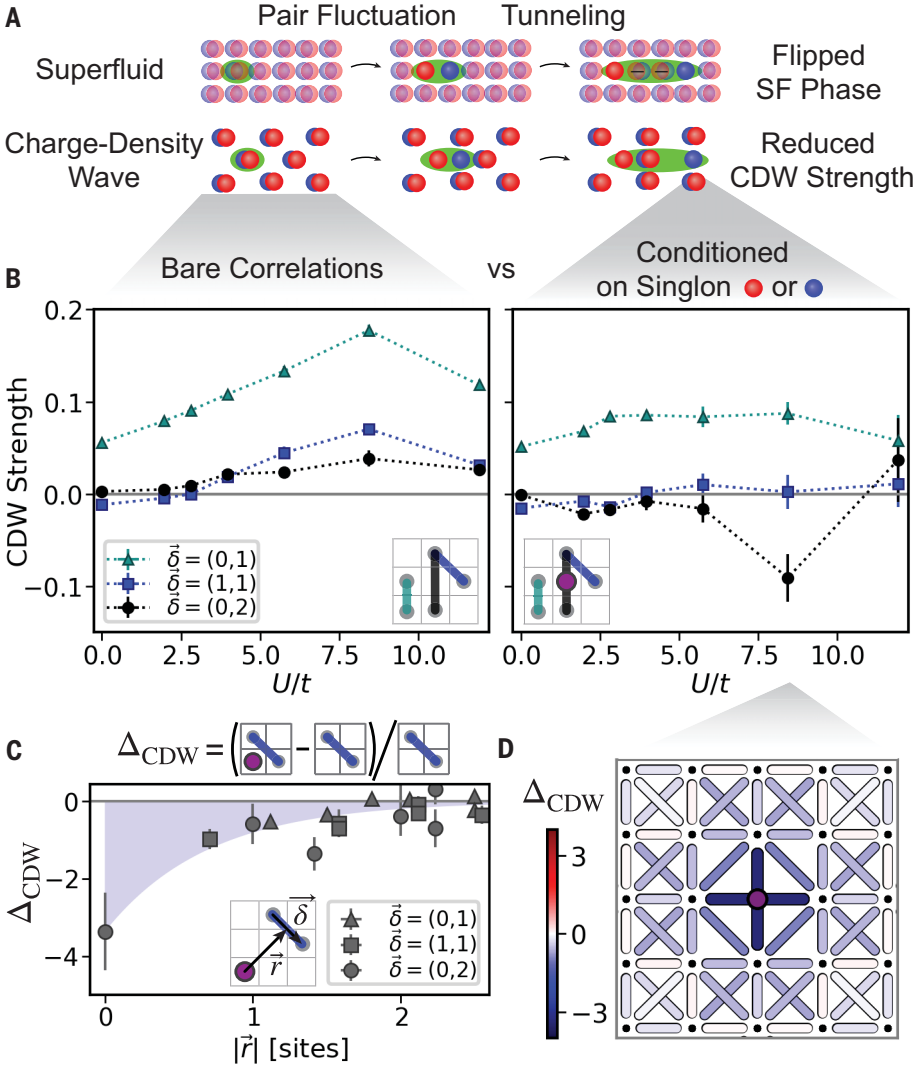



**Fig. 4. Interplay of nonlocal pairs and many-body order.** (**A**) When local pairs in a strongly correlated attractive system virtually fluctuate into spatially separated single spins, further tunneling can disturb the underlying many-body ordered state. (**B**) Polaronic correlations are observed by measuring disturbances of the CDW strength, defined as $\left\langle \left(\hat{d}_i - \hat{h}_i\right)\left(\hat{d}_{i+\delta} - \hat{h}_{i+\delta}\right)\right\rangle_c (-1)^{\delta_x+\delta_y}$, in the vicinity of a single spin. Whereas a peak in the background CDW strength is observed near $U/t \approx 8$ for various displacements $\vec{\delta}$ (left panel), a strong reduction in CDW strength is observed after conditioning on the presence of a nearby isolated spin (right panel). Right inset shows the relative location of the single spin (purple circle) and CDW bond. (**C**) The relative change in CDW strength $\Delta_{CDW}$ reveals the spatial extent of polaronic correlations at $U/t = 8.4(4)$. $\Delta_{CDW}$ decays with distance $|\vec{r}|$ from the singlon to the CDW bond. The shaded region represents the extent of the disturbance to the background order. (**D**) A symmetrized 2D map of (C), with lattice sites represented by black dots and the isolated spin at center. In the CDW strength, $\hat{d}_i$ ($\hat{h}_i$) denotes a doublon (hole, or empty site).

interaction (*23*), revealing that a single ↑ atom in total repels spin ↓ atoms on all other sites. This constitutes a strong direct signature of effective repulsion between pairs.

The development and destruction of CDW order across the phase diagram of Fig. 1 can be captured by the density response $\chi_n$ at wave vector $\vec{k} = (\pi, \pi)$ (Fig. 3C), which reflects the prevalence of low-energy states with checkerboard order. Highlighting the power of quantum gas microscopy, this thermodynamic property can be measured in equilibrium using the fluctuation-dissipation theorem for density perturbations, $\chi_n\left(\vec{k}\right) = (1/T)\sum_{\vec{\delta}} \langle \hat{n}_i \hat{n}_{i+\delta}\rangle_c \cos\left(\vec{k}\cdot\vec{\delta}\right)$, and the measured uniform density compressibility $\chi_n\left(\vec{k} = (0,0)\right) = \partial n/\partial \mu$ (Fig. 3D) (*25*). This same method allows measurement of $\chi_n\left(\vec{k}\right)$ throughout the Brillouin zone (Fig. 3C, inset) and provides a model-independent measurement of temperature $T$ (Fig. 3D) (*25*, *31*). The latter enables obtaining the magnetic susceptibility $\chi_m$ from the measured spin fluctuations without applying a magnetic field (*23*, *32*). As expected from the phase diagram in Fig. 1, the peak in CDW order occurs near $U/t \approx 6$. Also displayed are the interspin correlations, obtained from $\chi_{\uparrow\downarrow}\left(\vec{k}\right) = (1/T)\sum_{\vec{\delta}} \langle \hat{n}_{i\uparrow}\hat{n}_{i+\delta\downarrow}\rangle_c \cos\left(\vec{k}\cdot\vec{\delta}\right)$. Although opposite spins are uncorrelated at $U/t = 0$, they are seen to almost fully carry the CDW order beyond $U/t \approx 6$. Because density, magnetization, and interspin correlations are related by $\langle \hat{n}_i \hat{n}_{i+\delta}\rangle_c - \langle \hat{m}_i \hat{m}_{i+\delta}\rangle_c = 4\langle \hat{n}_{i\uparrow}\hat{n}_{i+\delta\downarrow}\rangle_c$, the relative agreement of $\chi_n$ and $\chi_{\uparrow\downarrow}$ illustrates the strength of density order as compared to magnetic order at $\vec{k} = (\pi, \pi)$. The pronounced CDW peak is also a signature of strong superfluid correlations within the crossover regime, as CDW correlations away from half-filling serve as a lower bound for superfluid correlations (*21*, *33*).

### Polaronic correlations

Given simultaneous charge and spin measurements, we finally explore the interplay of nonlocal pair fluctuations and the CDW order of other pairs, revealing the existence of polaronic correlations in the CDW order of the attractive Hubbard model. Polaronic correlations occur in the regime of highly nonlocal pairs, where further tunneling of a separated pair can dislocate the CDW checkerboard or flip the sign of superfluid correlations (Fig. 4A). These tunneling events prevent the virtual delocalization of other pairs across the bonds where the order has been reversed, costing an additional $4t^2/U$ per bond in the strong-coupling limit and further confining spatially separated pairs (*34*). This mechanism is directly complementary (*18*, *19*, *23*) to the magnetic polaron mechanism of the repulsive Fermi-Hubbard model (*35*, *36*), though here polaronic correlations dress the individual spins of a spatially separated fermion pair, rather than excess dopants.

In Fig. 4B, we compare the CDW correlations surrounding single spins to those present in the background. We quantify the underlying CDW strength as $\left\langle \left(\hat{d}_i - \hat{h}_i\right)\left(\hat{d}_{i+\delta} - \hat{h}_{i+\delta}\right)\right\rangle_c (-1)^{\delta_x+\delta_y}$ for a given displacement $\vec{\delta}$, which is positive for any $\vec{\delta}$ for a gas possessing checkerboard doublon-hole correlations. This underlying CDW strength peaks near an interaction strength $U/t \approx 8$. By contrast, for various $U/t$ values, the measured CDW strength is strongly reduced after conditioning on the presence of a single nearby isolated spin. This reduction substantially exceeds the lowest-order expectation of single-pair fluctuation events depicted in Fig. 4A, e.g., 25% for a displacement $\vec{\delta} = (0,1)$ and 50% for $\vec{\delta} = (1,1)$ or $(2,0)$. The measurements directly reveal the spatial extent of these polaronic effects, captured by the relative change $\Delta_{CDW}(|\vec{r}|)$ of conditioned to unconditioned CDW strength (Fig. 4, C and D). Virtual pair fluctuations disturb

the CDW order over a range of ~2 sites at $U/t = 8.4(4)$, with complete reduction or even reversal of the CDW order on nearby bonds. In future work, measurements of four-point correlations (*36*) around pairs of spins will further elucidate the internal structure of these quantum fluctuations.

## Conclusions

Our real-space observations of nonlocal fermion pairing and its interplay with CDW order illustrate the richness of the pseudo-gap regime of the attractive Hubbard model. Similar competing or intertwined orders are predicted for the repulsive Hubbard model. The methods can be extended further to study polaronic physics and superfluidity (*37*), pairing in momentum space as measured in bulk 2D gases (*38*), to detect the π phase shift of CDW order across stripes (*39*), and to directly measure the BCS condensate fraction through pair correlations (*40*).

## REFERENCES AND NOTES


1. M. Inguscio, W. Ketterle, C. Salomon, *Ultracold Fermi Gases* (IOS Press, 2008).
2. W. Zwerger, Ed., *The BCS-BEC Crossover and the Unitary Fermi Gas* (Springer, 2012).
3. M. Randeria, E. Taylor, *Annu. Rev. Condens. Matter Phys.* **5**, 209–232 (2014).
4. M. Randeria, N. Trivedi, A. Moreo, R. T. Scalettar, *Phys. Rev. Lett.* **69**, 2001–2004 (1992).
5. N. Trivedi, M. Randeria, *Phys. Rev. Lett.* **75**, 312–315 (1995).
6. P. A. Lee, N. Nagaosa, X.-G. Wen, *Rev. Mod. Phys.* **78**, 17–85 (2006).
7. Y. Cao *et al*., *Nature* **556**, 43–50 (2018).
8. Q. Chen, J. Stajic, S. Tan, K. Levin, *Phys. Rep.* **412**, 1–88 (2005).
9. R. T. Scalettar *et al*., *Phys. Rev. Lett.* **62**, 1407–1410 (1989).
10. J. Singer, T. Schneider, M. Pedersen, *Eur. Phys. J. B* **2**, 17–30 (1998).
11. T. Paiva, R. R. dos Santos, R. T. Scalettar, P. J. H. Denteneer, *Phys. Rev. B Condens. Matter Mater. Phys.* **69**, 184501 (2004).
12. T. Paiva, R. Scalettar, M. Randeria, N. Trivedi, *Phys. Rev. Lett.* **104**, 066406 (2010).
13. R. A. Fontenele, N. C. Costa, R. R. dos Santos, T. Paiva, *Phys. Rev. B* **105**, 184502 (2022).
14. J. M. Singer, M. H. Pedersen, T. Schneider, H. Beck, H. Matuttis, *Phys. Rev. B Condens. Matter* **54**, 1286–1301 (1996).
15. A. Moreo, D. J. Scalapino, *Phys. Rev. Lett.* **66**, 946–948 (1991).
16. M. Keller, W. Metzner, U. Schollwöck, *Phys. Rev. B Condens. Matter* **60**, 3499–3507 (1999).
17. J. Bauer, A. C. Hewson, N. Dupuis, *Phys. Rev. B Condens. Matter Mater. Phys.* **79**, 214518 (2009).
18. A. F. Ho, M. A. Cazalilla, T. Giamarchi, *Phys. Rev. A* **79**, 033620 (2009).
19. M. Gall, C. F. Chan, N. Wurz, M. Köhl, *Phys. Rev. Lett.* **124**, 010403 (2020).
20. P. T. Brown *et al*., *Nat. Phys.* **16**, 26–31 (2020).
21. D. Mitra *et al*., *Nat. Phys.* **14**, 173–177 (2018).
22. C. F. Chan, M. Gall, N. Wurz, M. Köhl, *Phys. Rev. Res.* **2**, 023210 (2020).
23. See supplementary materials.
24. J. Koepsell *et al*., *Phys. Rev. Lett.* **125**, 010403 (2020).
25. T. Hartke, B. Oreg, N. Jia, M. Zwierlein, *Phys. Rev. Lett.* **125**, 113601 (2020).
26. K. Yosida, *Phys. Rev.* **110**, 769–770 (1958).
27. M. Randeria, J.-M. Duan, L.-Y. Shieh, *Phys. Rev. B Condens. Matter* **41**, 327–343 (1990).
28. M. Drechsler, W. Zwerger, *Ann. Phys.* **504**, 15–23 (1992).
29. S. Murmann *et al*., *Phys. Rev. Lett.* **114**, 080402 (2015).
30. L. W. Cheuk *et al*., *Science* **353**, 1260–1264 (2016).
31. Q. Zhou, T.-L. Ho, *Phys. Rev. Lett.* **106**, 225301 (2011).
32. C. Sanner *et al*., *Phys. Rev. Lett.* **106**, 010402 (2011).
33. P. T. Brown *et al*., *Science* **357**, 1385–1388 (2017).
34. F. Grusdt *et al*., *Phys. Rev. X* **8**, 011046 (2018).
35. J. Koepsell *et al*., *Nature* **572**, 358–362 (2019).
36. J. Koepsell *et al*., *Science* **374**, 82–86 (2021).
37. J. K. Chin *et al*., *Nature* **443**, 961–964 (2006).
38. M. Holten *et al*., *Nature* **606**, 287–291 (2022).
39. T. Ying, R. T. Scalettar, R. Mondaini, *Phys. Rev. B* **105**, 115116 (2022).
40. W. Ketterle, M. Zwierlein, *Riv. Nuovo Cim.* **31**, 247–422 (2008).
41. T. Hartke, B. Oreg, C. Turnbaugh, N. Jia, M. Zwierlein, Replication Data for: Direct observation of non-local fermion pairing in an attractive Fermi-Hubbard gas, Dataverse (2023); https://doi.org/10.7910/DVN/HX07DM.
42. L. W. Cheuk *et al*., *Phys. Rev. Lett.* **114**, 193001 (2015).
43. T. Hartke, B. Oreg, N. Jia, M. Zwierlein, *Nature* **601**, 537–541 (2022).
44. R. Micnas, J. Ranninger, S. Robaszkiewicz, *Rev. Mod. Phys.* **62**, 113–171 (1990).
45. J. Bardeen, L. N. Cooper, J. R. Schrieffer, *Phys. Rev.* **108**, 1175–1204 (1957).
46. L. Salasnich, F. Toigo, *Phys. Rev. A* **86**, 023619 (2012).
47. M. Wortis, *Phys. Rev.* **132**, 85–97 (1963).
48. K. Miyake, *Prog. Theor. Phys.* **69**, 1794–1797 (1983).
49. C. N. Varney *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **80**, 075116 (2009).


## ACKNOWLEDGMENTS

**Funding:** This work was supported by the NSF through the Center for Ultracold Atoms and grant PHY-2012110, AFOSR (grant no. FA9550-16-1-0324 and MURI on molecules no. 2GG016303 PO No15323), and the Vannevar Bush Faculty Fellowship (ONR no. N00014-19-1-2631). **Authors contributions:** The experiment was designed by all authors. T.H., B.O., C.T., and N.J. collected and analyzed the data. All authors contributed to the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** The data and code for this manuscript are available on Dataverse (*41*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade4245
Supplementary Text
Figs. S1 to S9
References (*42*–*49*)

Submitted 15 August 2022; accepted 26 May 2023
10.1126/science.ade4245

BIOGEOGRAPHY

# Paleoenvironments shaped the exchange of terrestrial vertebrates across Wallace's Line

A. Skeels[1,2,3]*, L. M. Boschman[4], I. R. McFadden[1,2,5], E. M. Joyce[6], O. Hagen[7], O. Jiménez Robles[3,8], W. Bach[1,2], V. Boussange[1,2], T. Keggin[1,2], W. Jetz[9,10], L. Pellissier[1,2]

Faunal turnover in Indo-Australia across Wallace's Line is one of the most recognizable patterns in biogeography and has catalyzed debate about the role of evolutionary and geoclimatic history in biotic interchanges. Here, analysis of more than 20,000 vertebrate species with a model of geoclimate and biological diversification shows that broad precipitation tolerance and dispersal ability were key for exchange across the deep-time precipitation gradient spanning the region. Sundanian (Southeast Asian) lineages evolved in a climate similar to the humid "stepping stones" of Wallacea, facilitating colonization of the Sahulian (Australian) continental shelf. By contrast, Sahulian lineages predominantly evolved in drier conditions, hampering establishment in Sunda and shaping faunal distinctiveness. We demonstrate how the history of adaptation to past environmental conditions shapes asymmetrical colonization and global biogeographic structure.

The world's major biotic interchanges have had a disproportionate impact on the distribution of the world's terrestrial fauna (*1*–*3*). Well-known systems include the American interchange, which followed the formation of the Panamanian isthmus between North and South America (*4*), and the Indo-Australian interchange, which followed the convergence of the Australian and Eurasian tectonic plates and was shaped by complex dynamics of island formation (*5*). Far from an even diffusion of species between regions, interchanges have displayed an asymmetry in the direction of exchange (*6*–*8*) and have been dominated by particular functional traits (*9*, *10*). This unevenness may reflect a predisposition for dispersion in particular groups of organisms, because successful colonization is dependent on the ability both to disperse to a new region and to establish within new environmental conditions and ecological communities (*11*). Dispersal ability may act as a primary filter on colonization success (*12*), but lineages with broad environmental tolerances, or those with the ability to adapt to new environments, might contribute more to an interchange in the presence of environmental gradients between regions (*9*, *13*). To date, our ability to distinguish between hypotheses regarding the processes that shape the unevenness of different biotic interchanges has been limited by the available process-based methodologies and paleoenvironmental reconstructions.

Distributional patterns of terrestrial vertebrates in the Indo-Australian archipelago have fascinated naturalist for centuries (*14*) and have provided a model system for understanding biotic interchange. As the Australian continental plate approached the Eurasian plate, subduction at the northern boundary led to the formation of a geologically complex archipelago, today composed of thousands of islands and known biogeographically as Wallacea (*5*) (Fig. 1). The oceanic boundary separating the Sunda continental shelf (including Myanmar, Cambodia, Vietnam, Laos, Thailand, Malaysia, and Indonesia west of Lombok on the Eurasian plate) from Wallacea and the Sahul continental shelf (including Australia and New Guinea on the Australian plate) is named Wallace's Line after Alfred Russell Wallace (*14*) (Fig. 1), who noted a clear disjunction in the distribution of some taxa across it (*14*, *15*). Today, Wallace's Line marks the boundary of the Indomalayan and Australasian zoogeographic realms (*2*, *16*) (Fig. 1), but there are many lines in the region that mark turnover between a predominantly Sundanian or Sahulian fauna (*15*), such as the Heilprin-Lydekker Line (Fig. 1). The persistent presence of oceanic barriers has led to the dominant idea that dispersal limitation is the primary process shaping interchange (*12*). For example, the mammalian faunas of Sunda and Sahul are largely endemic, and only vagile rodents (*17*) and bats (*18*) have been successfully exchanged between continents. However, environmental niche traits may also have played a role (*19*, *20*) as the distinctly drier Australian continent (*20*, *21*) shaped the evolution of

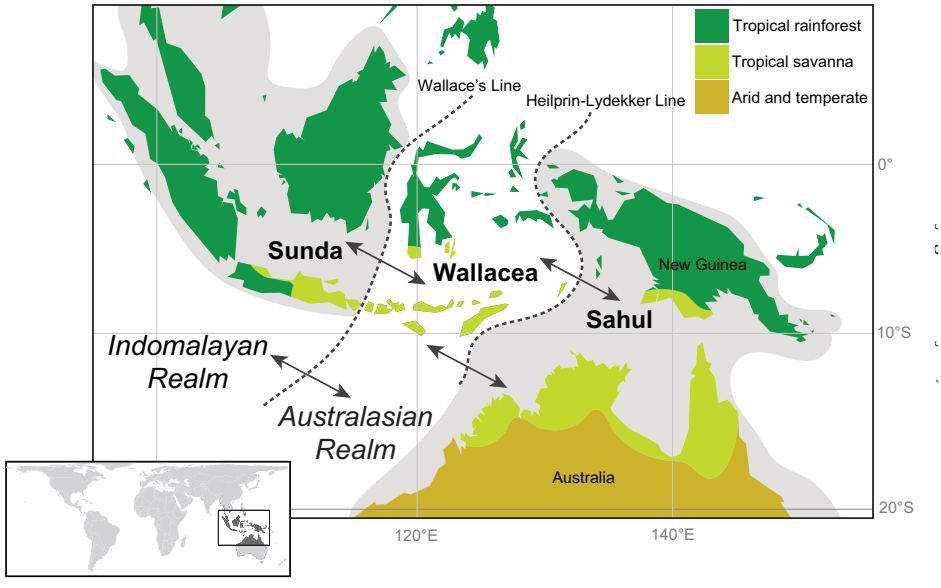



**Fig. 1. Climate and geography of the Indo-Australian archipelago at present.** The Sunda continental shelf is part of the Indomalayan zoogeographic realm and is separated from the Australasian zoogeographical realm by Wallace's Line. Wallacea is the island archipelago between the Sunda and Sahul continental shelves. The Sahul continental shelf, which includes the Australian mainland and New Guinea, is separated from Wallacea by the Heilprin-Lydekker Line. Sunda, Wallacea, and New Guinea are dominated by the tropical rainforest Köppen-Geiger climate belt, except for eastern Java, Bali, South Sulawesi, the Lesser Sunda isles, and southern New Guinea, which are dominated by the drier tropical savanna climate belt. The Australian mainland is dominated by tropical savanna, arid, and temperate climate belts. The climate maps are adapted from (*48*).

[1]Department of Environmental Systems Science, Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, ETH Zurich, 8092 Zurich, Switzerland. [2]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, 8903 Birmensdorf, Switzerland. [3]Research School of Biology, Australian National University, Canberra 0200, Australia. [4]Department of Earth Sciences, Utrecht University, 3584 CB Utrecht, Netherlands. [5]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE Amsterdam, Netherlands. [6]Systematics, Biodiversity and Evolution of Plants, Ludwig Maximilian University of Munich, 80331 Munich, Germany. [7]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany. [8]Institute of Biology, École Normale Supérieure, 75005 Paris, France. [9]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA. [10]Center for Biodiversity and Global Change, Yale University, New Haven, CT 06520, USA.
*Corresponding author. Email: alexander.skeels@gmail.com

an arid-adapted fauna (*22*). The interchange has been highly asymmetrical, with rates of exchange estimated to be at least twice as high from Sunda to Sahul than in the opposite direction (*23*). This could be shaped by the uneven sizes of the species pools, as stable areas of tropical rainforest in Sunda, connected to the Eurasian continent, supported a biodiverse fauna throughout the Cenozoic (*24*). Conversely, increasing aridity and a less connected continental area bound by oceans may have led to a reduced species pool in Sahul. Our current understanding of interchanges is pattern oriented, with historical biogeographic analyses determining where, when, and in which direction exchanges occur (*2*, *12*, *15*, *23*, *25*). A process-oriented understanding, however, requires the synthesis of paleogeographic and paleoclimatic information with biological traits and diversification histories across a range of taxa.

Paleoenvironmental and biological information united under a mechanistic modeling approach makes it possible to compare empirical observations with theoretical expectations and provides a framework with which to explore the mechanisms behind one of the world's major biotic interchanges. We hypothesize that, beyond a simple dispersal filter across oceanic barriers, paleoenvironmental niche dynamics also influence patterns of exchange. In addition to each clade's dispersal ability (*12*) and the size of the species pool, exchange patterns should be related to either the variability in the environmental conditions in which species occur (*25*) (mean realized niche breadth) or the rate at which environmental niche traits change over time (*13*) (rates of niche evolution). Further, the spatial distribution of climate, and how it has changed through time, should also explain asymmetry and the spatial distribution of lineages with either Sundanian or Sahulian ancestry (*19*). Here, we combined paleoenvironmental reconstructions of temperature, precipitation, and plate tectonics over the past 30 million years (Ma) (*26*–*28*) with a mechanistic model of biodiversity (*29*) to explore how dispersal and niche traits interact with the environment to shape regional diversification and patterns of exchange through time (*30*). The selected time period captures the onset of the convergence of the Australian and Eurasian plates (*5*) and the origin of most terrestrial vertebrate families in the region (fig. S1). We used published, well-sampled phylogenetic and spatial information on 20,433 species belonging to all 227 families of terrestrial vertebrates present in the Indo-Australian archipelago to estimate phylogenetic and taxonomic β diversity across Wallace's Line as measures of compositional dissimilarity between the two regions, as well as the number and direction of colonization events (*31*).

### Trait-based drivers of faunal exchange

We estimated a minimum of 381 independent colonization events across Wallace's Line based on a conservative phylogenetic hypothesis for each family [figs. S2 to S5 (*30*)], explaining the widespread distribution across Sunda and Sahul of 87 out of 227 terrestrial vertebrate families (Fig. 2 and Fig. 3A). Older colonization events were associated with greater species diversity (Spearman $\rho = 0.68$; Fig. 2A), most likely because these older lineages had a longer time to diversify. Precipitation niche breadth was the most consistent predictor of exchange in the Indo-Australian archipelago, indicating that colonization is more common in lineages that can tolerate wider variation in precipitation. Using phylogenetic multiple regression (table S1), we found that families with broader precipitation niche breadths were more likely to be exchanged between regions, to have a larger number of exchange events, and to have lower taxonomic β diversity and phylogenetic β diversity (Fig. 3B and table S1) across Wallace's Line. Temperature niche breadth was negatively associated with exchange dynamics (Fig. 3B and table S1), because temperature and precipitation niche breadths are inversely related in vertebrates (*32*). The

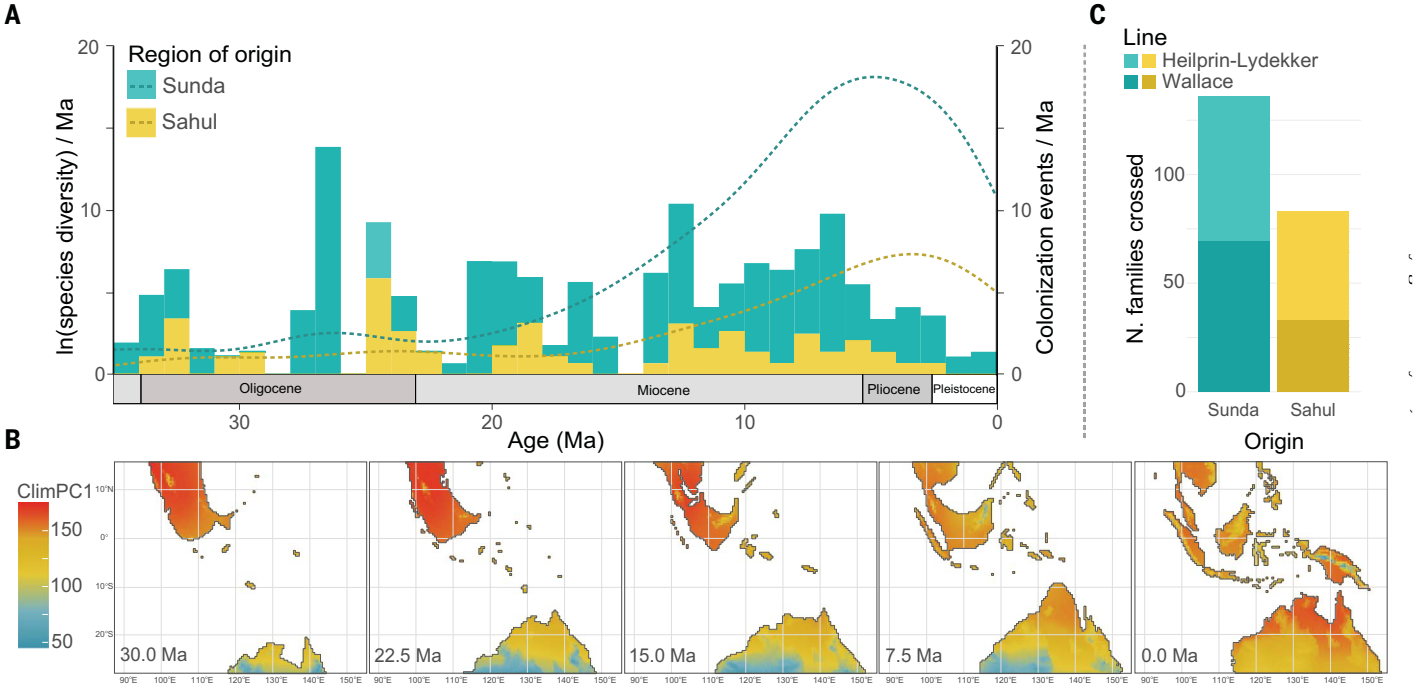



**Fig. 2. Colonization dynamics of terrestrial vertebrates and geoclimatic changes in Indo-Australia.** (**A**) Bars show the log-transformed number of species derived from independent colonization events across Wallace's Line from Sunda to Wallacea (blue) or in the opposite direction (yellow). Dashed lines show the number of colonization events through time. (**B**) Paleogeography (*28*) and paleoclimate (*26*, *27*) of the Indo-Australian archipelago over the past 30 Ma (*26*, *28*). (**C**) Numbers of families crossing between regions across two different biogeographic boundaries. Principle component analysis was performed on paleotemperature and paleoprecipitation to decompose a single climate axis (ClimPC1), explaining 88% of the variation in climate, where similar colors represent similar climates. Warm, wet conditions are shown in red; cold, dry conditions are shown in blue.

mechanistic model enabled the disentanglement of these factors and validated the role of broader precipitation niche breadths in the rate of exchange (table S2). Niche breadth may affect colonization dynamics through two pathways: (i) Clades more tolerant of a broad range of conditions may be less prone to extinction and therefore have a larger number of potential dispersing lineages (*8*), or (ii) they may be more able to establish in a diverse range of environmental conditions with greater establishment success. Our mechanistic model supports the second pathway, because higher rates of colonization were not associated with lower rates of extinction (fig. S6), but instead associated with certain dispersal and niche characteristics of the modeled lineages. Although the simulation results cannot exclude the role of differential extinction on colonization dynamics [e.g., (*8*)], they do suggest that differences in the rate of lineage diversification are not required to explain asymmetrical exchange. These results support the hypothesis that the precipitation gradient between Sunda and Sahul is a sharp barrier to the establishment of range-expanding lineages (*10*, *33*), and that lineages able to persist in a diversity of conditions can overcome this barrier.

Supplementing the effect of climatic niche breadth, our empirical and in silico model results support the importance of dispersal ability and the species pool in exchange patterns. More diverse biotas might facilitate interchange by having a greater number and variation of potential colonist species, increasing the number of potential colonization attempts (*34*), and this pathway is considered the null expectation in invasion biology (*35*). We found that the size of the source species pool was positively associated with exchange success and the number of colonization events in terrestrial vertebrates (Fig. 3B and table S1), as well as with all exchange metrics in the mechanistic model (table S2). As a key dispersal trait, flight was negatively associated with phylogenetic β diversity (Fig. 3B), and most exchanged taxa were highly volant, with 64% of families being either birds or bats,

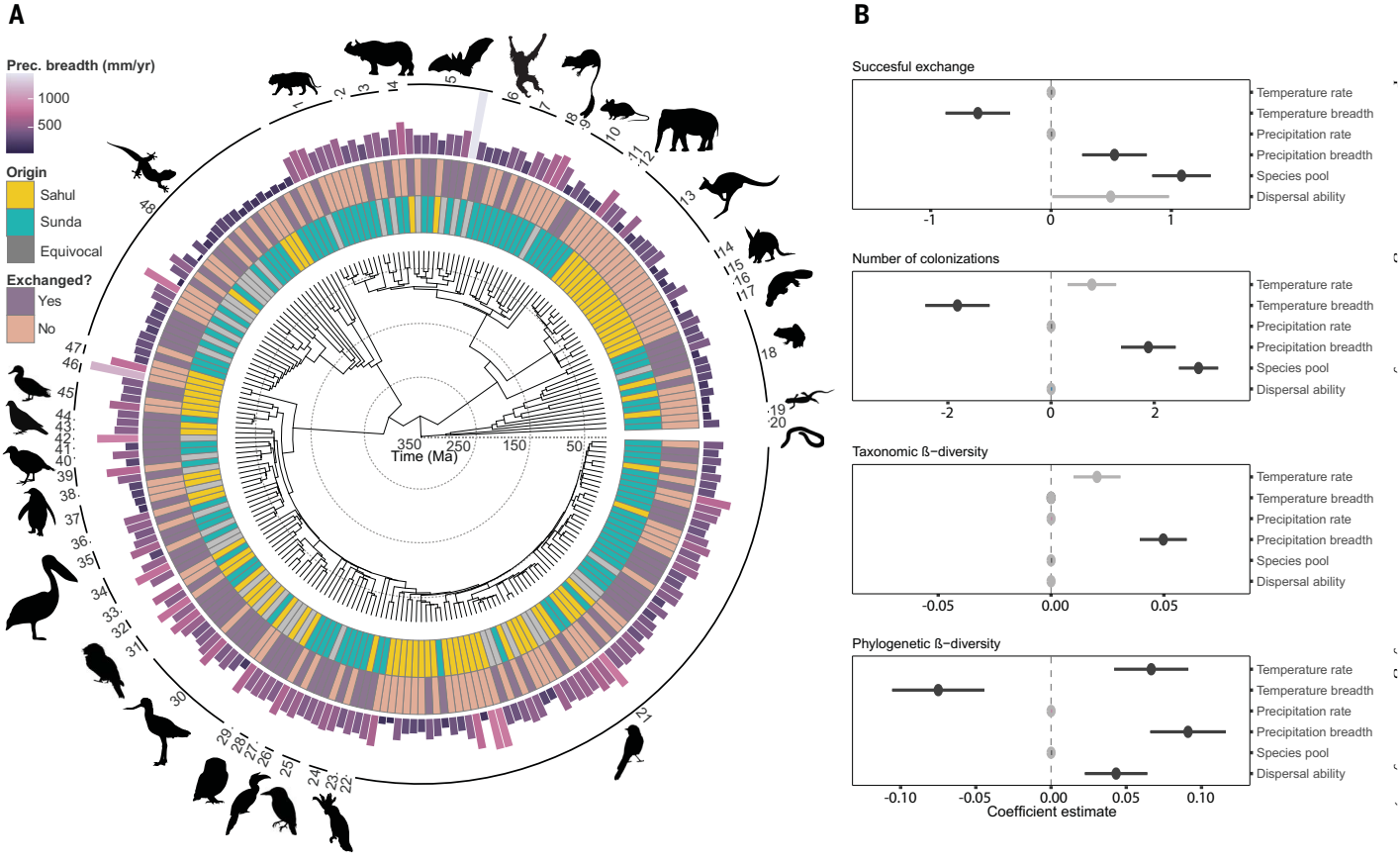



**Fig. 3. Exchange dynamics and precipitation niche breadth of terrestrial vertebrates on a family-level phylogenetic tree and effect size of the predictors of exchange.** (**A**) The circle in the center displays each family's reconstructed origin as Sunda, Sahul, or uncertain (equivocal support) based on biogeographic estimation models. The middle colored ring shows whether the family is present in both Sunda and Sahul and was reconstructed to have been exchanged across Wallace's Line. In the outer colored ring, the color and height of the bars represent the mean precipitation niche breadth of the family. Niche breadth was measured as the SD of mean annual rainfall across each species distribution (in millimeters per year). The numbered lines at the perimeter indicate the number of taxonomic orders, with representative taxa highlighted as silhouettes (images are from PhyloPic: http://phylopic.org/). 1, Carnivora; 2, Pholidota; 3, Cetartiodactyla; 4, Perrissodactyla; 5, Chiroptera; 6, Eulipotyphla; 7, Primates; 8, Scandentia; 9, Dermoptera; 10, Rodentia; 11, Lagomorpha; 12, Probosoidea; 13, Diprotodontia; 14, Dasyuromorphia; 15, Peramelemorphia; 16, Notoryctemorphia; 17, Monotremata; 18, Anura; 19, Caudata; 20, Gymnophiona; 21, Passeriformes; 22, Psittaciformes; 23, Falconiformes; 24, Piciformes; 25, Coraciiformes; 26, Bucerotiformes; 27, Trogoniformes; 28, Strigiformes; 29, Accipitriformes; 30, Charadriiformes; 31, Apodiformes; 32, Caprimulgiformes; 33, Eurypygiformes; 34, Suliformes; 35, Pelecaniformes; 36, Ciconiiformes; 37, Procellariiformes; 38, Sphenisciformes; 39, Gruiformes; 40, Cuculiformes; 41, Otidiformes; 42, Phaethontiformes; 43, Podicipediformes; 44, Columbiformes; 45, Galliformes; 46, Anseriformes; 47, Casuariiformes; 48, Squamata. (**B**) Coefficients of phylogenetic linear models of four exchange metrics with mean temperature and precipitation niche breadths, rates of temperature and precipitation niche evolution, dispersal ability, and species pool size. Predictors that were statistically significant based on phylogenetic linear models are in bold, and nonsignificant predictors are gray.

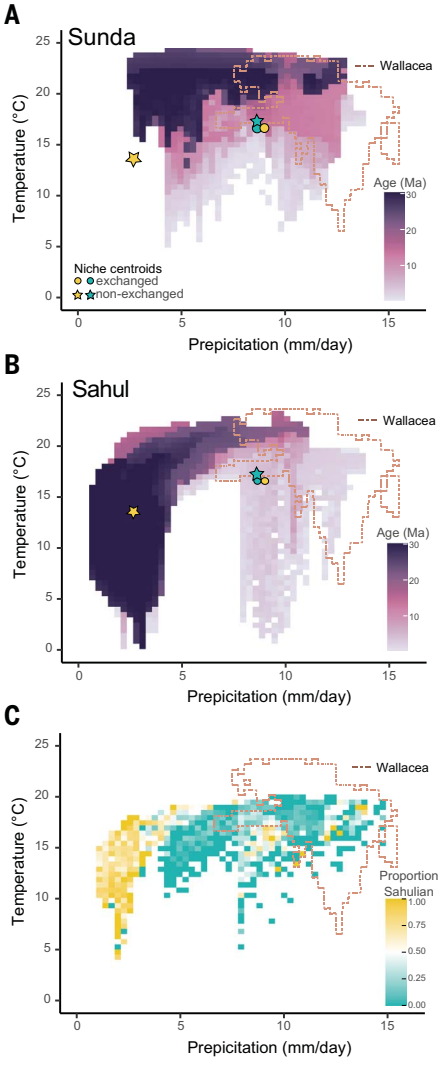



**Fig. 4. Age of paleoclimatic space in the Indo-Australian archipelago and distribution of climatic niche centroids of terrestrial vertebrates.** (**A** and **B**) Gridded representation of climate space defined by temperature and precipitation. Each tile in the grid represents a unique combination of these two variables. The tiles are colored by the age of that climate's first appearance in Sunda (A) and Sahul (B). Climate spaces that have been present longer are in darker colors. The climatic space occupied by the "stepping stone" region of Wallacea during the past 30 Ma is highlighted with an orange dashed line. This space has been present longer in Sunda than in Sahul. Symbols indicate the median climate niche centroid of species from Sahul (yellow) or Sunda (blue) that belong to families that have been exchanged (circle) or have not been exchanged (star). (**C**) Proportion of species of Sahulian origin in climate space. Species niche positions in climate space show that the area occupied by Wallacea is more densely filled with species from Sundanian families than with species from Sahulian families, and that species from both exchanged and nonexchanged Sundanian families are more commonly found in the climate space occupied by Wallacea.

accounting for 93% of the colonization events of Sunda. All exchange metrics were strongly positively associated with dispersal ability in the mechanistic model (table S2). We found that exchange of bird families across the Indo-Australian archipelago was proportionally higher than during other major biotic interchanges (*36*), reflecting the greater propensity of volant taxa to disperse over the deep-sea oceanic barriers such as Wallace's Line, which uniquely define the Indo-Australian interchange.

Contrary to our prediction, rates of niche evolution were not commonly associated with exchange dynamics (Fig. 3B and tables S1 and S2), possibly reflecting the widespread phenomenon of niche conservatism, the tendency for lineages to retain their ancestral environmental niche through time (*37*), during transoceanic colonization (*38*). For instance, many families that we reconstructed as having a Sundanian origin, such as Colubrid snakes and Pteropodid bats, are typically restricted to similar biomes in Sahul. Instead of the ability to rapidly adapt to new environmental conditions, we show that broader precipitation niche breadths and higher dispersal ability were the primary biological traits that allowed lineages to expand their geographic ranges in the face of oceanic barriers and a steep precipitation gradient.

### Paleoclimate and exchange asymmetry

We demonstrate that asymmetrical rates of exchange have been shaped by the distinctly different paleoclimatic histories of Sunda and Sahul. At least 46 families originating in Sunda (43% of all Sundanian families) crossed the Heilprin-Lydekker Line (Fig. 1 and table S3) to Sahul, whereas only 19 families originating in Sahul (24% of all Sahulian families) crossed Wallace's Line to Sunda. A necessary step to colonization involves crossing Wallacea (Fig. 1). Our paleoenvironmental reconstruction, which combines paleotopography (*28*), reconstructed Köppen belts (*27*), and a global circulation model (*26*), show that the humid climate space occupied by Wallacea emerged in Sunda at least 20 and 30 Ma (Fig. 4A and fig. S7) in a period associated with the spread of tropical rainforests (*21*, *24*). Conversely, in Sahul, areas with similarly humid conditions were present during the mid-Miocene (fig. S8) (*39*) and only became widespread in their contemporary location on Sahul at least 10 and 5 Ma (Fig. 4B), after the northward movement of Sahul and the uplift of New Guinea (*20*, *24*). An almost uninterrupted humid climate throughout the Indo-Australian archipelago, from Sunda to New Guinea, facilitated the establishment of Sundanian lineages in Sahul, and this can also explain the observation that Wallacean and New Guinean flora are compositionally more similar to the flora of Sunda than to that of the Australian mainland (*33*). Supporting this finding, we show that the climate space currently present in Wallacea and New Guinea is more densely occupied by vertebrate lineages from Sunda than by lineages from Sahul (Fig. 4C). By contrast, species from Sahul in families that did not colonize Sunda occupy an older and drier component of the Sahulian climate than species in families that colonized Sunda (Fig. 4C and fig. S9). A higher density of species in older climate space suggests that niche evolution has been truncated through time by the available climate on each continent, highlighting that niche conservatism limits exchange between Sunda and Sahul.

Greater connectivity of a humid climate at the interface of the exchange explains the observed patterns of asymmetrical interchange in terrestrial vertebrates, and this was supported by the mechanistic model. In the model, evolution of climatic niches was directed by the environmental conditions where species evolved (fig. S7). Therefore, Sundanian lineages evolved in niches closer to the conditions of Wallacea and colonized emergent Wallacean islands earlier (mean colonization = 15.27 Ma; fig. S10). In the opposite direction, Sahul lineages were generally unable to successfully colonize proto-Wallacea. Colonization events by Sahulian species across Wallace's Line generally occurred after the uplift of New Guinea's highlands (mean colonization time = 5.4 Ma; fig. S10), supporting the role of New Guinea as an ecological "stepping stone" between Wallacea and the Australian mainland (*40*).

We evaluated how climate connectivity explains exchange asymmetry by performing an in-silico counter-factual simulation experiment. Here, we either removed the precipitation or temperature constraints on species distributions or removed the tectonic history by holding the landscape constant with sea levels at the time period with maximum connectivity between the regions (Last Glacial Maximum, ~20,000 years ago). This effectively allowed dispersal within each continental region to be unconstrained by oceanic barriers and provided equivalent geographic isolation of each continental region from Wallacea. We found that the proportion of Sahulian species in Wallacea and New Guinea increased only when both the precipitation constraint and tectonic history were modified together, because this promoted more opportunities for dispersal and a higher likelihood of establishment (Fig. 5 and figs. S11 and S12). The proportion of simulations with colonization events between Sahul and Sunda was 16.6% in the full model, which was similar to the observed proportion in terrestrial vertebrates (24%). The proportion increased to 48.6% in the counterfactual model (fig. S11), which is more similar to the observed proportion of lineages moving in the opposite direction, from

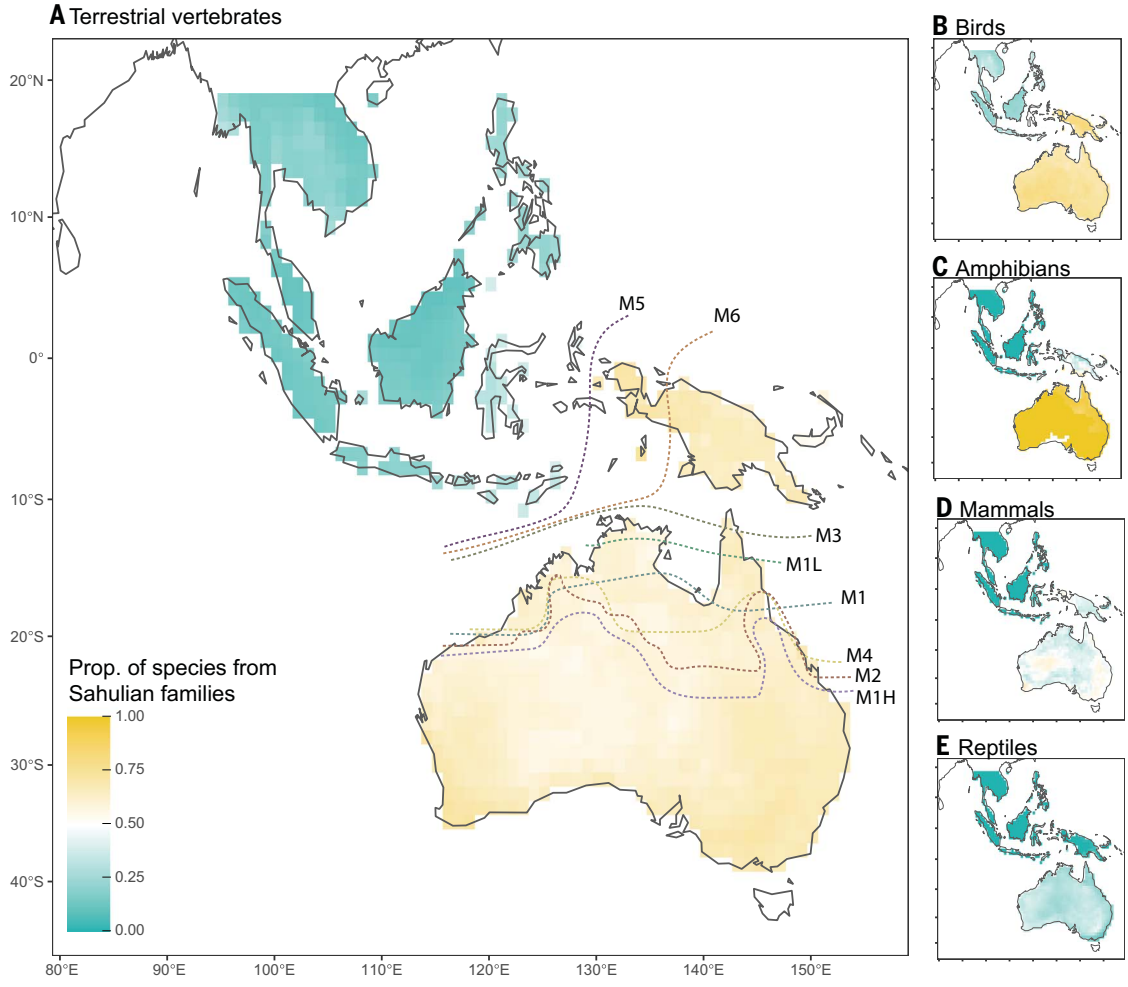



**Fig. 5. Spatial distribution of the proportion of terrestrial vertebrate species of Sahul origin and biogeographic turnover predictions from the biodiversity simulation model.** Proportion of total species richness in each grid cell that belongs to families of Sahul origin, based on biogeographic estimation, over the count of all species in the grid cell, for all terrestrial vertebrates (**A**), birds (**B**), amphibians (**C**), mammals (**D**), and reptiles (**E**). Dashed lines in (A) show the location of the sharpest spatial turnover of proportional Sahulian richness under six different simulations models (M1 to M6), as well as under high (M1H) and low (M1L) values of precipitation niche breadth and dispersal ability parameters of the full model (M1). The precipitation constraint on species distributions is removed in M2, the temperature constraint is removed in M3, tectonic history is removed in M4, both the precipitation constraint and tectonic history are removed in M5, and both the temperature constraint and tectonic history are removed in M6. The dashed lines illustrate how removing certain constraints that limit colonization changes the distribution of species from families with either Sahulian or Sundanian ancestry. Grid cell size is 110 × 110 km for all panels (Behrmann projection).

Sunda to Sahul (43%). Therefore, more symmetrical exchange could only be observed in the absence of climatic and dispersal barriers. A similar climate-mediated scenario has been suggested to explain exchange asymmetry during the American interchange (*9*, *41*) and between India and Eurasia (*7*, *42*). Our study clarifies how distinct climatic histories of continental regions can generate a near-ubiquitous pattern of asymmetrical terrestrial biotic interchanges.

### Dispersal and environmental niche shape spatial patterns of faunal exchange

Patterns of turnover among taxa have long been a source of contention for designating biogeographic boundaries in the Indo-Australian archipelago (*15*) because there is a high discordance in the distributional extent of different taxa across the Indo-Australian archipelago. Across the four classes of terrestrial vertebrates, gradients of turnover from a predominantly Sundanian to a predominantly Sahulian fauna differed, with a high proportion of sites in Sahul having high richness of Sundanian reptile species (100% of sites; Fig. 5E) and mammal species (68.8%; Fig. 5D) (which is driven by the exceptional diversity of Sundanian squamates, rodents, and bats), but not for amphibian species (11%; Fig. 5C) or bird species (1.4%; Fig. 5B). In the mechanistic model, a combination of dispersal and niche traits shaped gradients in biotic turnover, which could explain the emergence of these different patterns in vertebrates. When simulations were run with increased species' precipitation niche breadth and dispersal ability parameters, we found that lines demarcating turnover from a Sundanian to Sahulian biota (Fig. 5A) were located southward into regions of increased aridity, varying from a pattern corresponding with amphibians (Spearman's $\rho$ = 0.85 to 0.87; Fig. 5C) to a pattern corresponding with mammals ($\rho$ = 0.71 to 0.74; Fig. 5D), respectively. For birds, New Guinea is predominantly Sahulian, and this pattern was only reproduced by completely lifting the dispersal and precipitation constraints from the model

(ρ = 0.96; Fig. 5B and fig. S11). This could be because birds have the broadest precipitation tolerances and the highest volancy of any class of vertebrate, and therefore have fewer constraints on colonization (fig. S13). Because the simulation model is agnostic to the taxa investigated, we expect a similar mechanism to operate across a range of taxa not explored here, and there is some evidence that niche and dispersal traits are important drivers of interchange and patterns of turnover in plants (*10*, *43*).

## Conclusion

The longstanding view of biotic interchange is that it is primarily governed by plate tectonics and dispersal ability, with rates of colonization in terrestrial organisms being directly proportional to the source pool size (*1*), the geographic distance between regions, and the emergence of land bridges (*9*). A more complete picture emerges when the deep-time legacy of climate connectivity and niche conservatism on species distributions are also considered. Advances in deep-time climate reconstructions make it possible to test this directly, and support for a key role of climate connectivity is emerging in the Indo-Australian archipelago (*33*) and more broadly in other systems such as Indo-Eurasia (*7*, *42*) and the American interchange (*41*). The factors that shape colonization success are relevant in both a historical and a contemporary context, because biotic exchange is exacerbated by human intervention, and understanding these factors is therefore relevant for predicting the success of new biological colonizations and invasions (*35*). This is particularly true in the Indo-Australian archipelago, where biological ivasions are contributing to some of the highest rates of native species extinction on the planet (*45*). An emerging view of global biodiversity patterns is that they are largely the outcome of historical changes in plate tectonics shuffling lineages around the planet, in concert with major fluctuations in precipitation and temperature shaping dispersal, speciation, and extinction dynamics through environmental niches (*46*, *47*). A mechanistic modeling approach enables us to move beyond "lines" in biogeography and instead consider the processes that shape global variation in biodiversity patterns.

## REFERENCES AND NOTES


1. G. J. Vermeij, *Science* **253**, 1099–1104 (1991).
2. B. G. Holt *et al*., *Science* **339**, 74–78 (2013).
3. G. F. Ficetola, *Nat. Ecol. Evol.* **1**, 89 (2017).
4. F. G. Stehli, S. D. Webb, *The Great American Biotic Interchange* (Springer, 2013), vol. 4.
5. D. J. Lohman *et al*., *Annu. Rev. Ecol. Evol. Syst.* **42**, 205–226 (2011).
6. S. D. Webb, *Paleobiology* **17**, 266–280 (1991).
7. S. Klaus, R. J. Morley, M. Plath, Y. P. Zhang, J. T. Li, *Nat. Commun.* **7**, 12132 (2016).
8. J. D. Carrillo *et al*., *Proc. Natl. Acad. Sci. U. S. A.* **117**, 26281–26287 (2020).
9. J. T. Weir, E. Bermingham, D. Schluter, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21737–21742 (2009).
10. J. Y. S. Yap *et al*., *J. Biogeogr.* **45**, 838–847 (2018).
11. J. P. Sexton, P. J. McIntyre, A. L. Angert, K. J. Rice, *Annu. Rev. Ecol. Evol. Syst.* **40**, 415–436 (2009).
12. A. E. White, K. K. Dey, M. Stephens, T. D. Price, *Glob. Ecol. Biogeogr.* **30**, 685–696 (2021).
13. W. L. Eiserhardt, J. C. Svenning, W. J. Baker, T. L. P. Couvreur, H. Balslev, *Sci. Rep.* **3**, 1164 (2013).
14. A. R. Wallace, *Proc. R. Geogr. Soc. Lond.* **7**, 206–213 (1863).
15. J. R. Ali, L. R. Heaney, *Biol. Rev. Camb. Philos. Soc.* **96**, 922–942 (2021).
16. D. M. Olson *et al*., *Bioscience* **51**, 933–938 (2001).
17. K. C. Rowe *et al*., *J. Biogeogr.* **46**, 2752–2768 (2019).
18. S. M. Tsang *et al*., *J. Biogeogr.* **47**, 527–537 (2020).
19. R. Good, *Nature* **181**, 1763–1765 (1958).
20. H. A. Nix, in *Evolution of the Flora and Fauna of Arid Australia*, W. R. Barker, P. J. M. Greenslade, Eds. (Australian Systematic Botany Society, 1982), pp. 47–66.
21. R. J. Morley, *J. Trop. Ecol.* **34**, 209–234 (2018).
22. M. Byrne *et al*., *Mol. Ecol.* **17**, 4398–4417 (2008).
23. D. M. Crayn, C. Costion, M. G. Harrington, *J. Biogeogr.* **42**, 11–24 (2015).
24. R. Morley, in *Tropical Rainforest Responses to Climatic Changes*, M. Bush, J. Flenley, W. Gosling, Eds. (Springer, 2011), pp. 1–34.
25. K. A. Carscadden *et al*., *Q. Rev. Biol.* **95**, 179–214 (2020).
26. P. J. Valdes, C. R. Scotese, D. J. Lunt, *Clim. Past* **17**, 1483–1506 (2021).
27. C. R. Scotese, H. Song, B. J. W. W. Mills, D. G. van der Meer, *Earth Sci. Rev.* **215**, 103503 (2021).
28. R. Hall, *J. Limnol.* **72** (s2), e1 (2013).
29. O. Hagen *et al*., *PLOS Biol.* **19**, e3001340 (2021).
30. Materials and methods are available as supplementary materials.
31. N. J. Matzke, *Front. Biogeogr.* **5**, 242–248 (2013).
32. M. R. Pie, R. Divieso, F. S. Caron, *J. Zool. Syst. Evol. Res.* **59**, 1155–1166 (2021).
33. E. M. Joyce, K. R. Thiele, J. W. F. Slik, D. M. Crayn, *Biol. J. Linn. Soc. Lond.* **132**, 374–387 (2021).
34. J. L. Lockwood, P. Cassey, T. Blackburn, *Trends Ecol. Evol.* **20**, 223–228 (2005).
35. G. J. Vermeij, *Biol. Conserv.* **78**, 3–9 (1996).
36. C. D. Bacon *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6110–6115 (2015).
37. J. J. Wiens, *Evolution* **58**, 193–197 (2004).
38. M. D. Crisp *et al*., *Nature* **458**, 754–756 (2009).
39. K. J. Travouillon, S. Legendre, M. Archer, S. J. Hand, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **276**, 24–37 (2009).
40. E. Roycroft *et al*., *Curr. Biol.* **32**, 4215–4224.e3 (2022).
41. C. D. Bacon *et al*., *Geology* **44**, 375–378 (2016).
42. M. Bansal *et al*., *Science* **375**, 455–460 (2022).
43. L. A. Trethowan, *J. Veg. Sci.* **32**, 1–7 (2021).
44. S. C. Bernardes *et al*., *Sci. Rep.* **11**, 23357 (2021).
45. J. C. Z. Woinarski, A. A. Burbidge, P. L. Harrison, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4531–4540 (2015).
46. O. Hagen, A. Skeels, R. E. Onstein, W. Jetz, L. Pellissier, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026347118 (2021).
47. O. Hagen, *Ecography* **4**, e06132 (2022).
48. H. E. Beck *et al*., *Sci. Data* **5**, 180214 (2018).
49. A. Skeels *et al*., Data for: Paleoenvironments shaped the exchange of terrestrial vertebrates across Wallace's Line, Dryad (2023); https://doi.org/10.5061/dryad.1c59zw41f.


## ACKNOWLEDGMENTS

We thank the Ecosystems and Landscape Evolution Group at WSL and ETH Zurich for feedback, B. Flück for technical support, and M. Dawes for editing the manuscript for English. **Funding:** This work was supported by the Swiss National Science Foundation (grant 310030_188550 to L.P. and supporting A.S. and W.B.); ETH Zürich (internal grant ETH-34 18-1 to L.P.) and postdoctoral fellowship 18-2 FEL-52 to L.M.B.); the German Research Foundation (grant DFG-FZT 118, 202548816 to O.H.); the Swiss National Science Foundation (Postdoc Mobility Fellowship 206844 to I.M.); the European Union's Horizon 2020 Research and Innovation Programme (Marie Skłodowska-Curie grant 896323 to O.J.R.); and the Prinzessin Therese von Bayern Foundation (E.M.J.). **Author contributions:** Conceptualization: A.S., L.M.B., I.M., E.M.J., O.H., O.J.R., W.B., V.B., T.K., W.J., L.P.; Funding acquisition: L.P.; Methodology: A.S., L.M.B., I.M., O.H., V.B., T.K.; Investigation: A.S., L.M.B.; Project administration: A.S.; Supervision: L.P.; Visualization: A.S.; Writing – original draft: A.S.; Writing – review & editing: A.S., L.M.B., I.M., E.M.J., O.H., O.J.R., W.B., V.B., T.K., W.J., L.P. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data generated and analyzed during this study and the code required to replicate the analyses are available in the manuscript or have been deposited at Dryad (*49*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf7122
Materials and Methods
Figs. S1 to S13
Tables S1 to S3
References (*50*–*143*)

Submitted 9 November 2022; accepted 1 June 2023
10.1126/science.adf7122

MOLECULAR BIOLOGY

# Chromatin expansion microscopy reveals nanoscale organization of transcription and chromatin

Mark E. Pownall[1]*, Liyun Miao[1]*, Charles E. Vejnar[1], Ons M'Saad[2,3], Alice Sherrard[1], Megan A. Frederick[4], Maria D. J. Benitez[1], Curtis W. Boswell[1], Kenneth S. Zaret[4], Joerg Bewersdorf[2,3,5,6,7], Antonio J. Giraldez[1,8,9]*

Nanoscale chromatin organization regulates gene expression. Although chromatin is notably reprogrammed during zygotic genome activation (ZGA), the organization of chromatin regulatory factors during this universal process remains unclear. In this work, we developed chromatin expansion microscopy (ChromExM) to visualize chromatin, transcription, and transcription factors in vivo. ChromExM of embryos during ZGA revealed how the pioneer factor Nanog interacts with nucleosomes and RNA polymerase II (Pol II), providing direct visualization of transcriptional elongation as string-like nanostructures. Blocking elongation led to more Pol II particles clustered around Nanog, with Pol II stalled at promoters and Nanog-bound enhancers. This led to a new model termed "kiss and kick", in which enhancer–promoter contacts are transient and released by transcriptional elongation. Our results demonstrate that ChromExM is broadly applicable to study nanoscale nuclear organization.

Upon fertilization, embryos undergo transcriptional and cellular reprogramming to form a totipotent zygote (*1*–*5*). This reprogramming results in the transcriptional activation of genes required to initiate zygotic development. During this process, pioneer factors open the chromatin, recruit RNA polymerase II (Pol II), and activate transcription (*6*–*9*). However, the molecular organization of these factors during genome activation remains unclear. Nanoscale visualization of transcription and chromatin with current super-resolution approaches is limited by their ability to analyze nanoscale structure at the whole-nucleus scale, particularly in vivo. Alternatively, chromatin electron microscopy and tomography (ChromEMT) (*10*) provides nanometer resolution but lacks the multilabel imaging required to identify specific regulatory interactions. To overcome these limitations, we combined pan–expansion microscopy (pan-ExM) (*11*) with multimodal protein, RNA, and DNA labeling to resolve the nanoscale organization of chromatin, transcriptional activators, and transcription in a method that we term chromatin expansion microscopy (ChromExM). We applied ChromExM to uncover how the pioneer factor Nanog interacts with chromatin and recruits Pol II to activate transcription during genome reprogramming after fertilization.

[1]Department of Genetics, Yale University School of Medicine, New Haven, CT 06510, USA. [2]Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06510, USA. [3]Department of Biomedical Engineering, Yale University, New Haven, CT 06510, USA. [4]Institute for Regenerative Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [5]Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, CT 06510, USA. [6]Department of Physics, Yale University, New Haven, CT 06510, USA. [7]Nanobiology Institute, Yale University, West Haven, CT 06477, USA. [8]Yale Stem Cell Center, Yale University School of Medicine, New Haven, CT 06510, USA. [9]Yale Cancer Center, Yale University School of Medicine, New Haven, CT 06510, USA.
*Corresponding author. Email: mark.pownall@yale.edu (M.E.P.); liyun.miao@yale.edu (L.M.); antonio.giraldez@yale.edu (A.J.G.)

## Nanog forms DNA-bound foci associated with Pol II recruitment and transcription

In zebrafish, Nanog is required for genome reprogramming and transcriptional activation after fertilization (*8*, *9*, *12*, *13*). To investigate the molecular organization of transcriptional activation during genome activation, we analyzed the spatial organization of Nanog from the 32- to the 512-cell stage when transcription is activated. Live imaging of Nanog using a LlamaTag (*14*) revealed that Nanog formed clusters associated with the very early transcription of microRNA-430 (miR-430) (47% of miR-430 transcription sites contain Nanog foci) (Fig. 1, A and B; fig. S1, A to C) (*15*–*18*). Quantification of Nanog fluorescence revealed that these foci increased the local Nanog concentration 2.5- to 4-fold (fig. S1B), which is consistent with the occurrence of transcription factor and coactivator foci or hubs observed during embryogenesis and in cultured cells (*19*–*28*). We also observed that Nanog foci were associated with Pol II elongation, which is consistent with previous observations (*16*, *17*, *22*, *25*, *27*–*30*). Within a cell cycle, the formation of Nanog foci preceded Pol II elongation (Fig. 1, C and D), as visualized with a genetically encoded mintbody that detects elongating Pol II (Pol II pSer2) (*25*). After Pol II elongation begins, Nanog foci become less intense, which suggests that Nanog clusters may be evicted or dispersed after initiating transcription. In maternal-zygotic Nanog mutants (MZ*nanog*$^{-/-}$) (*8*), Pol II foci are lost and overall levels are reduced at the 256-cell stage, which confirms Nanog's role in Pol II recruitment and transcriptional activation (fig. S1, D and E) (*29*). Mutating the Nanog homeodomain to prevent DNA binding abolished Nanog foci, which indicates that they form in a DNA-binding–dependent manner rather than strictly through protein–protein interactions (fig. S1, F to I). Thus, Nanog forms DNA-bound hubs that increase its local concentration and are associated with Pol II recruitment and transcriptional activation during genome reprogramming.

## ChromExM achieves single-nucleosome resolution while preserving chromatin organization across scales

Next, we aimed to determine the underlying nanostructure and molecular organization of Nanog and RNA Pol II during chromatin reprogramming and transcriptional activation. To this end, we adapted the concept of pan-ExM (*11*) to include metabolic labeling of DNA and nascent RNA, along with antibody labeling to visualize the chromatin with nanometer resolution, hereafter termed ChromExM (Fig. 1, E and F). ChromExM involves the direct anchoring of biomolecules to a swellable hydrogel through the addition of acryloyl groups, which allows protein and nucleic acid retention during expansion (*11*). We achieved an average ~15× linear expansion factor of the nuclei in embryos (Fig. 1G), which corresponded to a ~4000-fold increase in nuclear volume and provided ~15-nm lateral resolution on a confocal microscope and ~3 nm in stimulated emission depletion (STED) super-resolution microscopy (Fig. 2).

Previous studies have reported conflicting results regarding the isotropy of ~4× to 8× expansion of chromatin (*31*, *32*). To address whether chromatin structure is perturbed by physical expansion, we developed an assay in which a pattern of parallel stripes is photocleaved into biotin-labeled DNA before expansion and visualized after expansion (Fig. 2A; fig. S2, A and B; and materials and methods). If chromatin expands isotropically, then the relative position of chromatin chains should be preserved, and the photocleaved stripes will remain parallel after expansion. We observed that stripes generated before expansion remain parallel after expansion and show no significant variation in their spacing when compared with simulated perfectly straight stripes generated as controls (Fig. 2, B and C; fig. S2, C and D; and materials and methods). This indicates that chromatin expands isotropically and the relative spatial organization of chromatin is preserved across the nucleus at submicron to global length scales by ChromExM (Fig. 2, B and C).

The preservation of chromatin structure at the nucleosomal level depends, in part, on the mesh size of the hydrogel, which affects the positional uncertainty and anchoring frequency

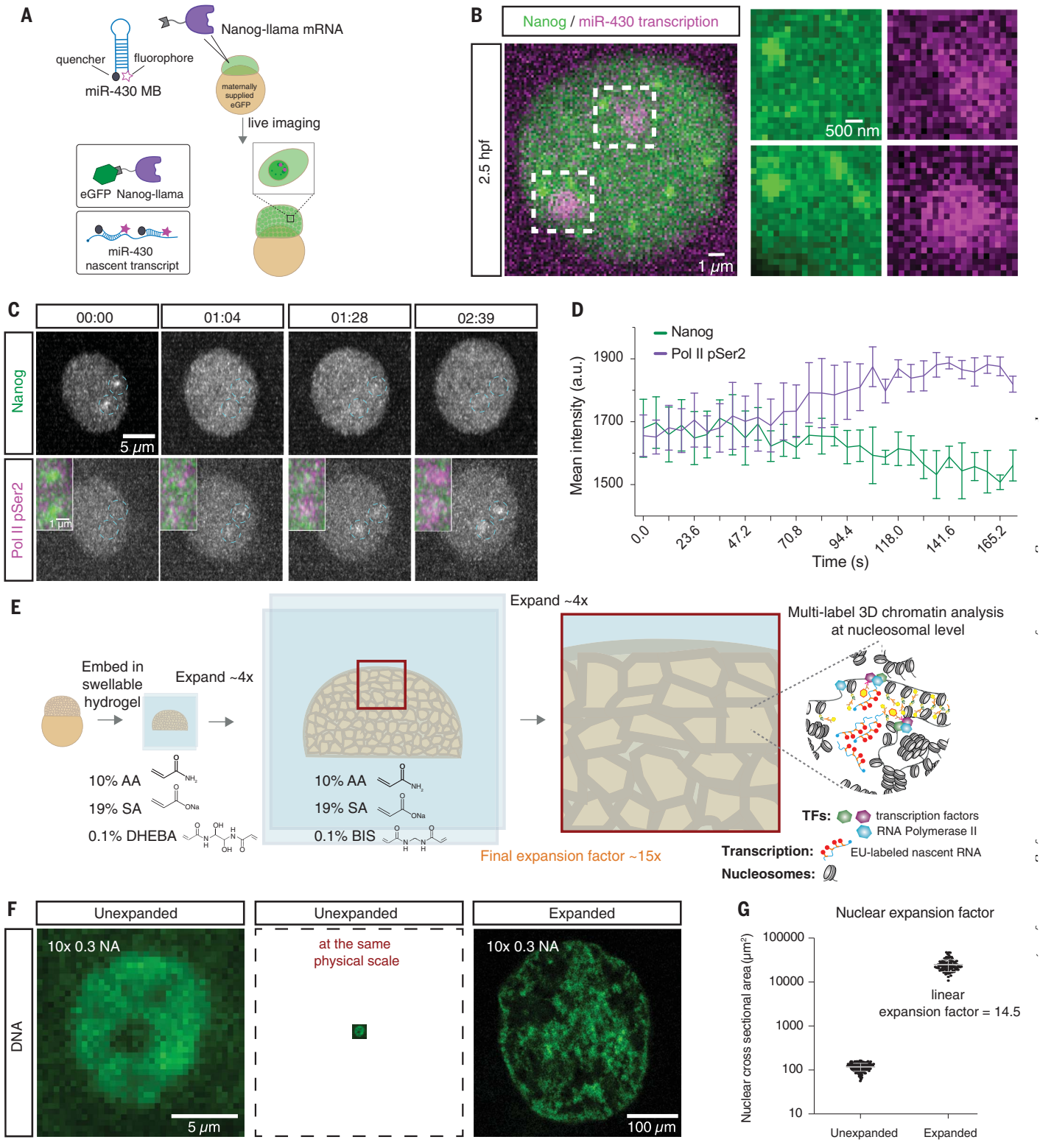



**Fig. 1. Nanog forms DNA-bound foci associated with zygotic transcription.** (**A**) Schematic that details how Nanog and miR-430 transcription were visualized in living embryos. eGFP, enhanced green fluorescent protein. (**B**) Live imaging shows Nanog foci associated with miR-430 transcription at 2.5 hpf. $n$ = 10 nuclei. (**C**) Live imaging shows Nanog foci are formed and disappear before Pol II elongation at = 2.75 hpf. $n$ = 2 nuclei, 1 embryo. (**D**) Quantification of fluorescence intensity in (C). a.u., arbitrary unit. (**E**) Schematic that shows the process of ChromExM and expected results. AA, acrylamide; BIS, N,N′-methylenebis(acrylamide); DHEBA, N,N′-(1,2-dihydroxyethylene) bisacrylamide; EU, 5-ethynyl uridine; SA, sodium acrylate. (**F**) Unexpanded and expanded nuclei from two unrelated embryos stained for DNA and imaged with a 10× 0.3 numerical aperture (NA) objective demonstrate the enhanced resolution provided by ChromExM. Scale bar is not corrected for the expansion factor. (**G**) Quantification of the nuclear expansion factor determined by measuring the nuclear cross-sectional area. $n$ = 171 unexpanded nuclei from 3 embryos and 104 expanded nuclei from 17 embryos.

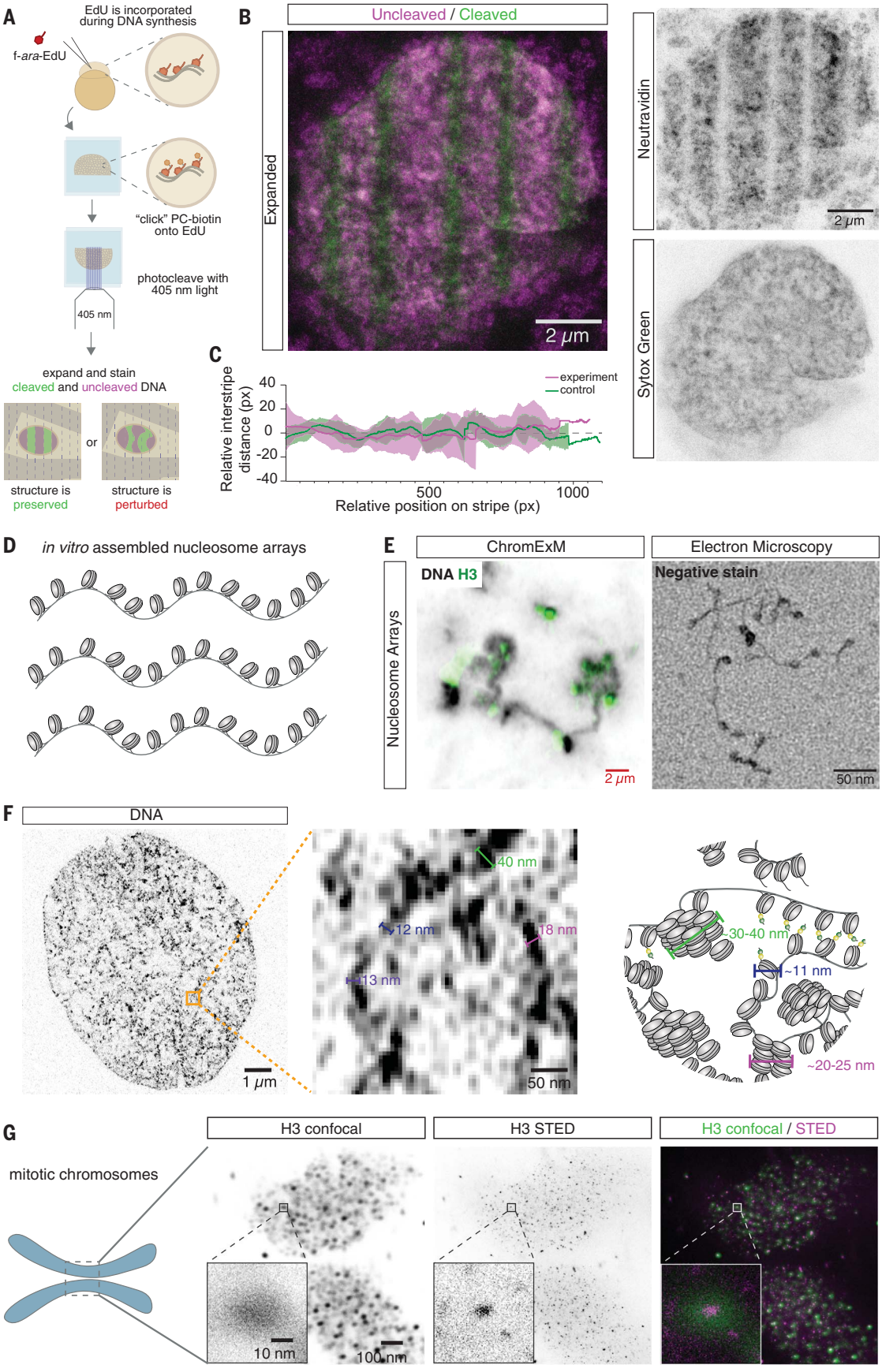



**Fig. 2. ChromExM preserves chromatin architecture and resolves chromatin fibers and individual nucleosomes.** (**A**) Schematic that shows how the chromatin is painted with photocleaved (PC) stripes to detect perturbations after expansion. (**B**) Expanded nucleus that shows that PC stripes remain parallel and sharp after expansion. The image is a maximum intensity projection of several *z*-slices. *n* = 20 nuclei from 3 embryos. (**C**) Quantification of relative interstripe distance in PC stripes versus a simulated control (methods) shows minimal variation in the spacing between stripes after expansion. *n* = 3 nuclei from 3 embryos. (**D**) Schematic of in vitro–assembled nucleosome arrays. (**E**) ChromExM image (left) that shows nucleosome arrays with H3 staining and EM image of nucleosome arrays with a similar conformation as the expanded array. Red scale bar is not corrected for the expansion factor. (**F**) ChromExM image with metabolic DNA labeling at 2.75 hpf that shows individual chromatin fibers and a schematic of chromatin. *n* = 6 nuclei, 5 embryos. (**G**) H3 staining imaged with confocal and STED microscopy resolves individual nucleosomes. *n* = 468 nucleosomes; 3 nuclei from 2 embryos.

of biomolecules to the gel (*33*). To determine whether chromatin structure is maintained at the nucleosome scale, we measured the mesh size of our swellable hydrogel by assessing the mobility of differently sized molecules in the gel using fluorescence recovery after photobleaching (fig. S3, A to E) (*34*). We observed that a 3000 molecular weight (MW) dextran recovered quickly after photobleaching with a diffusion coefficient of 0.72 $\mu m^2/s$, but histone H1 and a 2,000,000 MW dextran did not recover (fig. S3, A to E). These results demonstrate that the mesh size is small enough to prevent rapid diffusion of the ~32-kDa H1 protein, which suggests that the gel polymer can anchor the chromatin with subnucleosomal resolution. To functionally test this, we compared ChromExM and electron microscopy (EM) of in vitro–assembled nucleosome arrays containing 13 nucleosomes along 2.7 kb of DNA (Fig. 2, D and E; fig. S3, F and G) (*35*). We observed a similar organization of nucleosome arrays in ChromExM and unexpanded EM, with an average of 12.7 and 12.5 nucleosomes detected per array, respectively (Fig. 2E and fig. S3, F to H). Taken together, these results demonstrate that the local chromatin organization can be maintained at the nucleosomal scale during ChromExM.

To visualize chromatin, we metabolically labeled the DNA with (2′S)-2′-deoxy-2′-fluoro-5-ethynyluridine (f-*ara*-EdU) (*36*), followed by fluorescent picolyl azide detection after expansion (*37*). This approach was more photostable than the intercalating dye SYTOX Green and increased the labeling intensity ~45-fold compared with standard azides (fig. S3, J to L), which improved chromatin labeling to overcome the effects of molecular decrowding and reduced brightness caused by the ~4,000-fold volumetric expansion. This approach resolved chromatin fibers in the nucleus (Fig. 2F) with similar diameters (<12 to 40 nm) to those determined by EM (*10*). Nucleosomes are ~10 nm in size, which is at the resolution limit of confocal microscopy after ChromExM. Therefore, we performed STED microscopy with ChromExM to improve the resolution by an additional factor of approximately five, which approached <3-nm lateral resolution, and demonstrated that ChromExM can resolve individual nucleosomes by using histone H3 immunostaining and confocal microscopy (~80% of nucleosomes detected by confocal are individual nucleosomes by STED) (Fig. 2G and fig. S3I). The combination of ChromExM with STED provides fast multimodal molecular imaging of the chromatin that approaches <3-nm resolution.

### Visualizing interactions between nucleosomes and the pioneer factor Nanog

We used ChromExM to determine the underlying nanoscale organization of transcription hubs during genome activation by visualizing the molecular-scale interactions among the pioneer factor Nanog (*8*, *9*), nucleosomes, and Pol II at 4 hours post fertilization (hpf). To analyze how Nanog interacts with chromatin, we costained for Nanog and H3 and visualized an average of 88,619 and 104,347 individual particles of Nanog and H3 per nucleus, respectively, with a false-detection rate of 0.05% for identifying Nanog particles (Fig. 3, A to D; fig. S4, A to G; and movie S1). Quantitative analysis of the spatial organization of Nanog and H3 particles revealed several classes of Nanog–nucleosome conformations (Fig. 3, B to G; fig. S4, A to E and H to I; and materials and methods). In the first class, Nanog was closely associated with nucleosomes (<20 nm distance between Nanog and H3), with 2.0% of H3 and 3.3% of Nanog particles falling into this group (Fig. 3, B and E, and fig. S4A), which likely indicates a bound state in which Nanog is potentially initiating chromatin opening. In the second class, a subset of class 1, Nanog clusters were bound to a nucleosome (>1 Nanog particle within 20 nm and additional Nanog particles within 50 nm), which represents recruitment of multiple Nanog particles to the same chromatin region (0.06% of H3, 0.2% of Nanog) (Fig. 3, C and F, and fig. S4, B and B′). The third class contained Nanog particles distantly associated with nucleosomes (20 to 100 nm away from each other) (26.9% of H3, 39.7% of Nanog), which potentially indicates chromatin regions already opened by Nanog (Fig. 3, D and G, and fig. S4C). The fourth class contained nucleosomes and Nanog not closely associated with each other (>100 nm away) (71.2% of H3, 57.0% of Nanog) (fig. S4, D and E). Our analysis identified 88,619 Nanog particles per nucleus, which is consistent with the number of chromatin immunoprecipitation followed by sequencing (ChIP-seq) peaks (~40,000) identified at a similar stage in zebrafish (*8*, *38*). These measurements are likely the lower bounds of Nanog–nucleosome interactions considering that some nucleosomes may not be detected here. Taken together, we are able to characterize pioneer factor–nucleosome organization during chromatin opening.

### RNA Pol II forms string nanostructures associated with nascent RNA

To understand how transcription is organized during chromatin reprogramming, we used ChromExM to visualize Nanog and Pol II phosphoserine 5 (pSer5), which is deposited on Pol II after recruitment to the promoter (Fig. 3, H to K; fig. S5, A to C; and movie S2) (*39*–*41*). We detected an average of 27,503 Pol II particles per nucleus, which have characteristics consistent with single-molecule detection on the basis of their size and homogeneous intensity (fig. S5, D to G). Globally, the average distance between Pol II pSer5 and Nanog particles was 94 nm (Fig. 3L), with 19% of Pol II pSer5 particles being within 50 nm of Nanog particles, consistent with Nanog's role in Pol II recruitment (*8*). However, we observed that Pol II exhibited three distinct types of organization (Fig. 3, I to K, and fig. S5, A, B, C, and H). The first class involved large groupings of interspersed Pol II pSer5 and Nanog particles (412 Pol II pSer5 particles and 100 Nanog particles on average, 132-μm mean length) arranged like beads on a string, which we refer to as class 1 strings and occurred twice per nucleus in 86% of nuclei observed (Fig. 3, I and M to O; fig. S5A). RNA fluorescence in situ hybridization (RNA FISH) for miR-430 revealed that class 1 strings represent sites of active miR-430 transcription (Fig. 3, P and Q; fig. S6, A to C). To validate this result, we assembled the highly repetitive miR-430 locus using single haploid embryo long-read genome sequencing and identified a single, continuous ~550-kb miR-430 locus (fig. S7, A to F), which corresponded to an estimated length of ~200 μm, which is consistent with the 132-μm length of class 1 strings (Fig. 3O). This locus has been estimated to have >300 promoters capable of transcribing primary miR-430 transcripts (*42*), which is consistent with the ~1800 mature miR-430 genes detected in our assembly (fig. S7, C and D). Similarly, we detect ~400 Pol II pSer5 particles in class 1 strings, which likely represent binding at these promoters (Fig. 3M). The string-like structure that connects Pol II and Nanog particles is consistent with Pol II and Nanog binding profiles observed by ChIP-seq at this locus (fig. S7, G and H) (*8*, *12*). Additionally, miR-430 RNA FISH signal and class 1 strings are lost in *miR-430*$^{-/-}$ embryos (figs. S6, A and B, and S8, A and B) (*43*), which we confirmed lack all miR-430 genes using long-read genome sequencing (fig. S7, C to F). Together, these results demonstrate that class 1 strings show the nanoscale organization of the transcriptionally active miR-430 cluster (*16*, *17*).

The second class of Pol II pSer5 organization involved multiple Pol II particles organized into strings with an average length of 831 nm (2 to 70 particles; mean of 4 particles per string) (Fig. 3, J and M to O; fig. S5, B and H). In this class, 37% of Pol II pSer5 particles were within 50 nm of Nanog particles, consistent with Nanog's role in Pol II recruitment. We hypothesized that these linear class 2 Pol II pSer5 strings represent individual genes loaded with Pol II arranged in single-file lines. Class 2 strings are still present in *miR-430*$^{-/-}$ embryos (fig. S8, B and C), which indicates that they represent other transcribed genes (*8*). This processive organization of Pol II is reminiscent of active transcriptional elongation. In other cases, we observed branched Pol II pSer5 strings (Fig. 3J and fig. S5, B and H), which is consistent with the

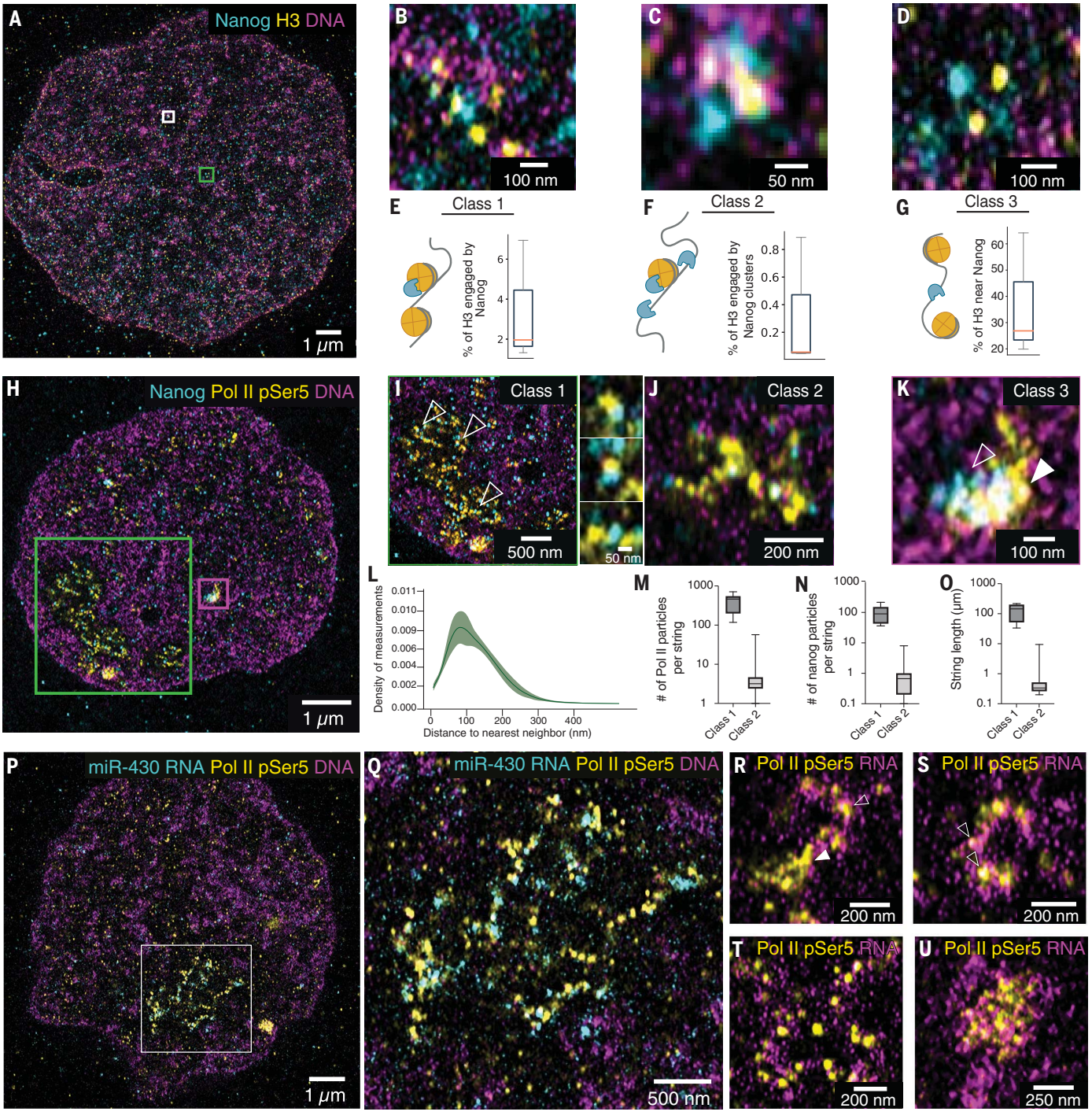



**Fig. 3. Visualizing the nanoscale organization of Nanog, nucleosomes, Pol II, and transcription.** (**A** to **D**) Representative images showing Nanog and H3 staining at 4 hpf; (B) represents class 1 organization. (C) represents the white boxed region in (A) and shows class 2 organization. (D) represents the green boxed region in (A) and shows class 3 organization. *n* = 3 nuclei from 2 embryos. (**E** to **G**) Box plots quantifying the percentage of H3 particles in each class of Nanog–H3 interactions. The median is denoted by the orange line. *n* = 104,347 distances. (**H** to **K**) Representative images showing Nanog and Pol II pSer5 staining at 4 hpf, *n* = 7 nuclei from 2 embryos. (I) represents the green boxed region in (H) and shows class 1 organization; areas indicated with arrowheads are magnified to the right. (J) represents class 2 organization. (K) represents the magenta boxed region in (H) and shows class 3 organization; arrowheads indicate regions enriched for Nanog (open arrowhead) and Pol II pSer5 (closed arrowhead). (**L**) Density plot of the distance to nearest neighbor for Nanog and Pol II pSer5 particles; *n* = 169,866 distances. (**M**) Box plots quantifying the number of Pol II pSer5 particles in class 1 (*n* = 36) and class 2 (*n* = 5682) strings. (**N**) Box plots quantifying the number of Nanog particles in class 1 and class 2 Pol II pSer5 strings. (**O**) Box plots quantifying the length of Pol II pSer5 strings in class 1 and class 2. (**P** and **Q**) ChromExM images showing Pol II pSer5 and miR-430 RNA detected by hybridization chain reaction RNA FISH at 4 hpf. (Q) represents the boxed region in (P). (**R** to **U**) Representative images showing nascent RNA and Pol II pSer5 at 4 hpf. *n* = 7 nuclei from 3 embryos. The closed arrowhead in (R) indicates a central Pol II pSer5 hub with an emanating string of individual Pol II pSer5 particles associated with nascent transcripts (open arrowhead). Open arrowheads in (S) indicate individual Pol II pSer5 particles.

transcription factory model in which multiple genes are in close proximity and share a regulatory Pol II hub (*29*, *44*, *45*). Indeed, we detected an average of 822 Pol II pSer5 strings per nucleus, which is consistent with the ~1700 active zygotic genes at 4 hpf we previously identified (*17*).

The third class of Pol II organization involved "macroclusters" (Fig. 3K and fig. S5C) similar to those previously observed by super-resolution imaging (*20*–*22*). We detected an average of 28 Pol II pSer5 macroclusters per nucleus. In cases in which these macroclusters were occupied by both Nanog and Pol II, we were able to resolve distinct regions within the cluster occupied solely by Nanog and Pol II (Fig. 3K and fig. S5C), as if the two factors were tethered on distinct DNA elements. These examples are consistent with Nanog bound to enhancers that reside in close proximity (<50 nm) to Pol II bound at the promoter (*8*, *38*). Similarly, we observe all three classes of Pol II pSer5 organization during the minor wave of genome activation at 2.5 hpf, although there were fewer active sites, which is consistent with lower levels of transcription at this time (fig. S9, A to H) (*16*, *17*, *46*). This suggests that each class of Pol II organization represents a generalizable state of Pol II organization broadly used throughout genome activation. At 4 hpf, we observed that there are an average of 5.3 Pol II particles within 200 nm of each Nanog particle among all three classes, which suggests that multiple Pol IIs are recruited by Nanog. Among the macroclusters, we also observed multiple cases in which a string of RNA Pol II emanated from the cluster (Fig. 3K), which raises the question of how these two distinct Pol II pSer5 structures may function.

We hypothesized that Pol II pSer5 strings represent actively transcribing Pol II, which has exited the shared regulatory hub where Nanog, Pol II, and potentially other coactivators congregate to control gene expression. To test this, we combined metabolic RNA labeling using 5-ethynyl uridine with ChromExM labeling for Pol II pSer5 and observed nascent transcripts associated with all three classes of Pol II pSer5 structures (Fig. 3, R to U; fig. S10, A to F; and movie S3). We observed a wide range of transcription levels associated with the Pol II pSer5 strings, which suggests that they may encompass actively transcribed gene bodies as well as paused promoters (fig. S10G). The organization of these transcriptionally active Pol II pSer5 strings is consistent with Pol II elongation along the gene body of individual genes, for which we visualized multiple Pol II pSer5 particles that extrude nascent transcripts in concert (Fig. 3, R to T; fig. S10, A, B, C, E, and F). Taken together, these results support a model in which Nanog and Pol II clusters are in close physical proximity and form a regulatory hub that activates transcription as Pol II exits the hub to form a string along the gene body (Fig. 3R and fig. S10A).

### Enhancers and promoters are kicked apart by transcription elongation

Next, we assessed whether these Pol II pSer5 structures are formed in a transcription-dependent manner (Fig. 4, A to H; fig. S11, A to G; and movie S4). Inhibiting transcription elongation with α-amanitin (*47*–*50*) led to a strong reduction of Pol II pSer5 string length (66% reduction; $P = 0.002$) (Fig. 4, C, D, G, H, and I) and a mild reduction in the size and number of Pol II macroclusters (26.3% reduction in number; $P = 0.0304$) (Fig. 4J and fig. S11H) without reducing Pol II particle count (fig. S11I). Taken together, these data indicate that Pol II pSer5 strings represent transcription elongation, whereas Pol II pSer5 macroclusters may function as a regulatory hub formed independently of elongation.

We also observed that Nanog and Pol II particles were closer after treatment with α-amanitin (Fig. 4K; median distance 106 nm versus 66 nm; $P < 0.001$), with more Pol II pSer5 particles surrounding each Nanog particle (Fig. 4L; 3.1 versus 7.6 Pol II per Nanog; $P < 0.001$). Similar behavior between Nanog and Pol II was observed when transcription was inhibited in *miR-430*$^{-/-}$ embryos (fig. S11J), which suggests that these changes occur at multiple loci in the genome. We explain these results with the following model: Nanog first recruits Pol II to enhancer–promoter hubs, which brings them into close proximity; then Pol II exits the regulatory hub in the form of Pol II pSer5 strings, which transcribes the gene body (Fig. 4). When transcription elongation was inhibited, we observed an increase in the stoichiometry and proximity of Nanog-Pol II particles, which is consistent with Pol II stalling at the promoter and the stabilization of enhancer and promoter interactions.

To test this model, we performed ChIP-seq for Pol II pSer5 in wild-type and α-amanitin–treated embryos and observed extensive pileup of Pol II at promoters when blocking transcription (Fig. 4M and fig. S11K). As predicted from our ChromExM results, Pol II pSer5 accumulated at Nanog-bound enhancers and other accessible sites (Fig. 4, N and O; fig. S11, L and M), yet these regions are not enriched for RNA Pol II in wild-type embryos (Fig. 4, N and O; fig. S11, L and M) (*8*). These results can be explained by a model in which enhancer–promoter contacts are transient, consistent with the kiss-and-run model (*21*), and there is, therefore, a larger mean distance between Nanog and Pol II (Fig. 4K). However, in the absence of transcriptional elongation, Pol II is continuously brought to the promoter and remains in close proximity to the enhancer as observed by ChIP-seq. This is further supported by the increase in the number of Pol II particles that surround each Nanog particle (Fig. 4L). On the basis of these data, we propose a modified version of the kiss-and-run model termed "kiss and kick," in which transcription itself kicks away the enhancer from the promoter as Pol II elongates (Fig. 4P). This would explain why we observe a stabilization of Poll II in close proximity to Nanog-bound enhancers when elongation is inhibited.

### Discussion

Here, we developed ChromExM for multimodal super-resolution chromatin imaging by physically enlarging biological samples to achieve nucleosomal resolution. We demonstrate that chromatin organization is preserved and developed an improved metabolic labeling strategy for chromatin imaging (Fig. 2). ChromExM provides markedly higher resolution than previous ExM applications for chromatin imaging (~3 to 15 nm versus ~65 nm) (*31*). Other methods to visualize chromatin, such as ChromEMT (*10*), lack multimodal labeling, and the resolution of single-molecule localization microscopy is limited by the size of the fluorescent labels (~20 nm for primary and secondary antibody), which becomes negligible in ChromExM given that labels are applied after expansion (*11*). Although ChromExM provides substantial technical advances, it requires bright and photostable fluorescent labeling and chemical fixation, which may affect the appearance of certain structures. Future methods will be needed to identify specific loci through DNA FISH compatible with ChromExM to investigate specific regulatory structures at individual genes.

We used ChromExM to characterize the nanoscale organization of the pioneer factor Nanog and RNA Pol II as they activate transcription during genome activation. Although previous studies have shown how pioneer factors direct chromatin opening in vitro (*51*, *52*) and organize into hubs in vivo (*19*, *28*, *53*, *54*), the nanoscale organization of such hubs has remained unclear. ChromExM revealed that Pol II shows three types of organization during genome activation and is intimately associated with Nanog (Fig. 3, H to K). Previous studies have shown that transcription factors (TFs) such as Oct4, Brd4, and Mediator form clusters associated with superenhancers (*21*, *23*, *24*). With the resolution of ChromExM, we can now resolve how Nanog and Pol II often occupy distinct regions within transcription hubs, which is consistent with their binding at enhancers and promoters that are in close contact.

How enhancers and promoters are organized to activate transcription is central to understanding gene regulation. Previous reports have concluded that enhancer–promoter contact is correlated (*55*), anticorrelated (*56*), or unrelated (*57*) with transcription, which

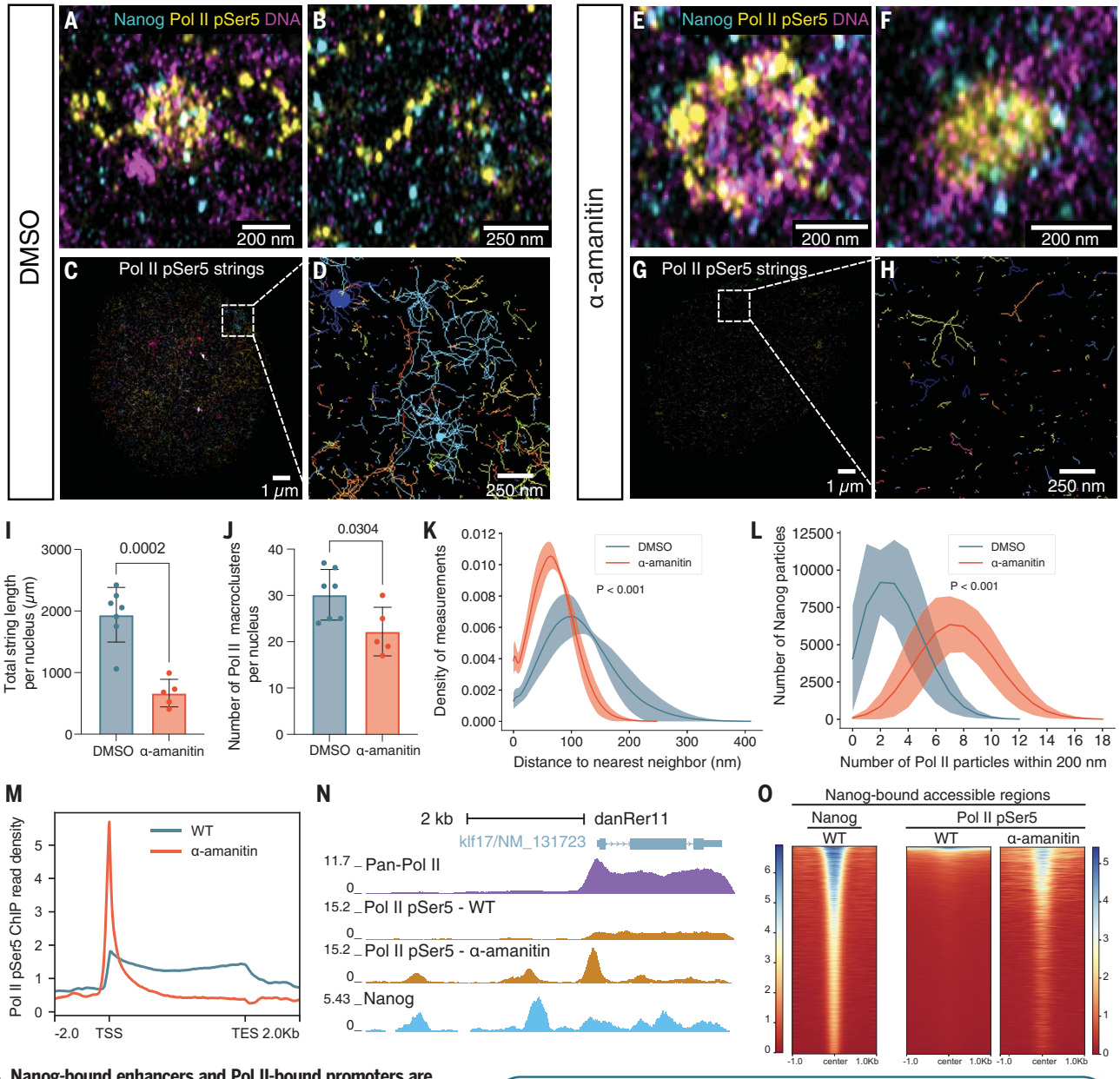



**Fig. 4. Nanog-bound enhancers and Pol II-bound promoters are kicked apart by transcription elongation.** (**A** and **B**) Representative images of Nanog and Pol II pSer5 in dimethyl sulfoxide (DMSO)–treated embryos at 4 hpf. $n$ = 6 nuclei from 2 embryos. (**C** and **D**) Visualization of Pol II pSer5 strings identified in DMSO-treated embryos at 4 hpf. (**E** and **F**) Representative images of Nanog and Pol II pSer5 in α-amanitin–treated embryos. (**G** and **H**) Visualization of Pol II pSer5 strings identified in α-amanitin–treated embryos. (**I**) Quantification of total Pol II pSer5 string length per nucleus in DMSO- and α-amanitin–treated embryos. $P$ = 0.0002; unpaired $t$ test. (**J**) Quantification of the number of Pol II pSer5 macroclusters detected in DMSO- and α-amanitin–treated embryos. $P$ = 0.0304; unpaired $t$ test. (**K**) Density plot of the distance to nearest neighbor for Nanog to Pol II pSer5 particles in DMSO- and α-amanitin–treated embryos. $n$ = 221,335 and 229,706 distances, respectively. $P$ < 0.001; Mann-Whitney $U$ test. (**L**) Histogram showing the number of Pol II pSer5 particles within 200 nm of each Nanog particle in DMSO- and α-amanitin–treated embryos. $P$ < 0.001; Mann-Whitney $U$ test. (**M**) Line plot showing Pol II pSer5 binding across gene bodies ±2 kb at zygotic genes in wild-type (WT) and α-amanitin–treated embryos. TES, transcription end site TSS, transcription start site. (**N**) Representative genome tracks of pan-Pol II, Pol II pSer5, and Nanog binding showing accumulation of Pol II pSer5 at the promoter and Nanog-bound enhancers in the presence of α-amanitin. (**O**) Heatmaps showing Pol II pSer5 binding at Nanog-bound accessible regions in WT and α-amanitin–treated embryos. Regions are ranked by Nanog ChIP-seq signal. (**P**) Schematic showing the kiss-and-kick model.

leaves the question of how enhancer–promoter contact is related to transcription unanswered. Three models have been proposed to explain how enhancers and promoters interact to control gene expression: (i) stable contact between enhancers and promoters, (ii) a dynamic kissing model with transient enhancer–promoter contacts, and (iii) a TF activity gradient in which TFs diffuse from enhancers to promoters rather than requiring physical contact (*58*). Our results are not consistent with models (i) and (iii) and lead to a model in which Nanog can form clusters at enhancers, coming into physical proximity with Pol II at the promoter and eventually triggering transcription, which we observed outside of this shared regulatory hub as strings of Pol II. Consistent with this model, we found that Pol II strings were substantially reduced when elongation was inhibited, which coincided with an accumulation of Pol II at promoters and in Nanog-bound enhancers and indicates that the enhancer and promoter were stabilized in close contact (Fig. 4, A to P). Taken together, these results suggest that Pol II elongation displaces enhancer–promoter contacts and lead us to propose the kiss-and-kick model. In this model, the dynamic association between the enhancer and promoter is kicked apart during elongation. This effect can be caused by either transcription elongation or the nascent RNAs that have been shown to dissolve Mediator condensates (*59*). The kiss-and-kick model could also explain transcriptional bursting (*60*), as elongation would be triggered in intervals while the enhancer and promoter are in contact and then paused after elongation kicks away the enhancer.

## REFERENCES AND NOTES


1. N. L. Vastenhouw, W. X. Cao, H. D. Lipshitz, *Development* **146**, dev161471 (2019).
2. M. T. Lee, A. R. Bonneau, A. J. Giraldez, *Annu. Rev. Cell Dev. Biol.* **30**, 581–613 (2014).
3. V. Yartseva, A. J. Giraldez, *Curr Top Dev Biol.* **113**, 191–232 (2015).
4. K. N. Schulz, M. M. Harrison, *Nat. Rev. Genet.* **20**, 221–234 (2019).
5. C. L. Wike *et al*., *Genome Res.* **31**, 981–994 (2021).
6. A. Soufi, G. Donahue, K. S. Zaret, *Cell* **151**, 994–1004 (2012).
7. A. Soufi *et al*., *Cell* **161**, 555–568 (2015).
8. L. Miao *et al*., *Mol. Cell* **82**, 986–1002.e9 (2022).
9. M. Pálfy, G. Schulze, E. Valen, N. L. Vastenhouw, *PLOS Genet.* **16**, e1008546 (2020).
10. H. D. Ou *et al*., *Science* **357**, eaag0025 (2017).
11. O. M'Saad, J. Bewersdorf, *Nat. Commun.* **11**, 3850 (2020).
12. M. T. Lee *et al*., *Nature* **503**, 360–364 (2013).
13. M. Leichsenring, J. Maes, R. Mössner, W. Driever, D. Onichtchouk, *Science* **341**, 1005–1009 (2013).
14. J. P. Bothma, M. R. Norstad, S. Alamos, H. G. Garcia, *Cell* **173**, 1810–1822.e16 (2018).
15. P. Heyn *et al*., *Cell Rep.* **6**, 285–292 (2014).
16. Y. Hadzhiev *et al*., *Nat. Commun.* **10**, 691 (2019).
17. S. H. Chan *et al*., *Dev. Cell* **49**, 867–881.e8 (2019).
18. A. J. Giraldez *et al*., *Science* **312**, 75–79 (2006).
19. M. Mir *et al*., *eLife* **7**, e40497 (2018).
20. W.-K. Cho *et al*., *eLife* **5**, e13617 (2016).
21. W.-K. Cho *et al*., *Science* **361**, 412–415 (2018).
22. I. I. Cisse *et al*., *Science* **341**, 664–667 (2013).
23. B. R. Sabari *et al*., *Science* **361**, eaar3958 (2018).
24. A. Boija *et al*., *Cell* **175**, 1842–1855.e16 (2018).
25. S. Uchino *et al*., *J. Cell Biol.* **221**, e202104134 (2022).
26. Y. Sato *et al*., *Development* **146**, dev179127 (2019).
27. S.-K. Huang, P. H. Whitney, S. Dutta, S. Y. Shvartsman, C. A. Rushlow, *Curr. Biol.* **31**, 5102–5110.e5 (2021).
28. K. Kuznetsova *et al*., *Curr. Biol.* **33**, 164–173.e5 (2023).
29. F. J. Iborra, A. Pombo, D. A. Jackson, P. R. Cook, *J. Cell Sci.* **109**, 1427–1436 (1996).
30. A. Pombo *et al*., *EMBO J.* **18**, 2241–2253 (1999).
31. E. L. Faulkner *et al*., *J. Cell Sci.* **135**, jcs259009 (2022).
32. S. P. Pernal *et al*., *Histochem. Cell Biol.* **153**, 469–480 (2020).
33. F. Chen, P. W. Tillberg, E. S. Boyden, *Science* **347**, 543–548 (2015).
34. J. Karvinen, T. O. Ihalainen, M. T. Calejo, I. Jönkkäri, M. Kellomäki, *Mater. Sci. Eng. C* **94**, 1056–1066 (2019).
35. M. A. Frederick *et al*., *Nat. Struct. Mol. Biol.* **30**, 31–37 (2023).
36. A. B. Neef, N. W. Luedtke, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20404–20409 (2011).
37. C. Uttamapinant *et al*., *Angew. Chem. Int. Ed.* **51**, 5852–5856 (2012).
38. C. Xu *et al*., *Dev. Cell* **22**, 625–638 (2012).
39. S. Buratowski, *Mol. Cell* **36**, 541–546 (2009).
40. H. Lu, L. Zawel, L. Fisher, J.-M. Egly, D. Reinberg, *Nature* **358**, 641–645 (1992).
41. S. Trigon *et al*., *J. Biol. Chem.* **273**, 6769–6775 (1998).
42. Y. Hadzhiev *et al*., *Dev. Cell* **58**, 155–170.e8 (2023).
43. Y. Liu *et al*., *PLOS ONE* **8**, e76387 (2013).
44. J. A. Mitchell, P. Fraser, *Genes Dev.* **22**, 20–25 (2008).
45. C. S. Osborne *et al*., *Nat. Genet.* **36**, 1065–1071 (2004).
46. M. Ugolini, K. Kuznetsova, H. Oda, H. Kimura, N. L. Vastenhouw, Targeted disruption of transcription bodies causes widespread activation of transcription. bioRxiv 517317 [Preprint]. 21 November 2022.
47. D. A. Bushnell, P. Cramer, R. D. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1218–1222 (2002).
48. A. L. Greenleaf, *J. Biol. Chem.* **258**, 13403–13406 (1983).
49. T. J. Lindell, F. Weinberg, P. W. Morris, R. G. Roeder, W. J. Rutter, *Science* **170**, 447–449 (1970).
50. D. A. Kane *et al*., *Development* **123**, 47–55 (1996).
51. M. Fernandez Garcia *et al*., *Mol. Cell* **75**, 921–932.e6 (2019).
52. E. Luzete-Monteiro, K. S. Zaret, *Curr. Opin. Struct. Biol.* **75**, 102425 (2022).
53. J. Dufourt *et al*., *Nat. Commun.* **9**, 5194 (2018).
54. M. M. Gaskill, T. J. Gibson, E. D. Larson, M. M. Harrison, *eLife* **10**, e66668 (2021).
55. H. Chen *et al*., *Nat. Genet.* **50**, 1296–1303 (2018).
56. N. S. Benabdallah *et al*., *Mol. Cell* **76**, 473–484.e7 (2019).
57. L. J. Mateo *et al*., *Nature* **568**, 49–54 (2019).
58. J. P. Karr, J. J. Ferrie, R. Tjian, X. Darzacq, *Genes Dev.* **36**, 7–16 (2022).
59. J. E. Henninger *et al*., *Cell* **184**, 207–225.e24 (2021).
60. A. Coulon, C. C. Chow, R. H. Singer, D. R. Larson, *Nat. Rev. Genet.* **14**, 572–584 (2013).
61. M. E. Pownall, C. E. Vejnar, A. J. Giraldez, Code for: Chromatin expansion microscopy reveals nanoscale organization of transcription and chromatin (1.0), Zenodo (2023); https://doi.org/10.5281/zenodo.7933370.


## ACKNOWLEDGMENTS

We thank C. Castaldi, I. Tikhonova, G. Wang, and E. Ng from the Yale Center for Genome Analysis for sequencing support; the Yale Center for Genome Analysis and Keck Microarray Shared Resource at Yale University for providing the necessary PacBio sequencing services, which is funded in part by the National Institutes of Health instrument grant 1S10OD028669-01; S. Dube, T. Gerson, M. Messih, A. Bonneau, and J. P. Fernandez for technical help; P. Kidd and Y. Tian for assistance with STED microscopy; J. Bothma and H. Garcia for assistance with Llama tag constructs; the Yale Center for Advanced Microscopy Facility for the use of their SP8, LSM 880, and FEI Biotwin microscopes; the Marine Biological Laboratory at Woods Hole and B. Sözen for use of equipment; L. Lavis for sharing Janelia Fluor dyes; G. Riddihough, S. Hiley, and I. Carmi for feedback on the manuscript; and all members of the Giraldez lab for feedback and support. This work is supported by the 4DNucleome program. **Funding:** National Institutes of Health grant F31HD10444 (M.E.P.), National Institutes of Health grant T32HD007149 (M.E.P.), Yale Genetics Venture Fund and the Surdna foundation (L.M.), National Institutes of Health grant F31DK123886 (M.A.F.), National Institutes of Health grant T32DK07780 (M.A.F.), National Institutes of Health grant T32GM0077223 (M.D.J.B.), Canadian Institute of Health Research (C.W.B.), National Institutes of Health grant GM36477 (K.S.Z.), Wellcome Trust grant 203285/B/16/Z (J.B.), National Institutes of Health grant P30 DK045735 (J.B.), National Institutes of Health grant R01 HD100035 (A.J.G.), and National Institutes of Health grant R35 GM122580 (A.J.G.). **Author contributions:** M.E.P. and A.J.G. conceived of the project. M.E.P. developed ChromExM, performed all light microscopy, developed the image analysis pipelines, analyzed all imaging data, and prepared samples for PacBio sequencing; L.M. performed ChIP-seq, analyzed the data, and generated haploid embryos. C.E.V. developed python scripts for distance analysis and analyzed the PacBio data; O.M. assisted with the initial development of ChromExM; A.S. performed EM; M.A.F. prepared nucleosome arrays; M.E.P., L.M., M.D.J.B., and C.W.B. generated reagents for imaging; K.S.Z., J.B., and A.J.G. supervised the project and acquired funding; M.E.P. and A.J.G. wrote the manuscript; M.E.P., L.M., C.E.V., A.S., M.A.F., K.S.Z., J.B., and A.J.G. revised the manuscript. **Competing interests:** M.E.P and A.J.G are inventors on a provisional patent application (no. 63/500,193) filed by Yale University with the US patent office covering the ChromExM method described here. J.B. has financial interests in Bruker Corp. and Hamamatsu Photonics. O.M. and J.B. filed patent applications (no. US-20230132184-A1) with the US patent office covering the pan-ExM method. O.M. and J.B. are co-founders of panluminate, Inc., which is developing related products. The other authors declare no competing financial interests. **Data and materials availability:** Information on high-throughput sequencing data is available in data S1. Raw high-throughput sequence reads generated in this study are publicly available in the Sequence Read Archive under project accession no. SRP434969/PRJNA962791. Scripts used in this study are available at Zenodo (*61*). Other materials generated in this study are available upon request.
**License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade5308
Materials and Methods
Figs. S1 to S11
References (*62–101*)
Data S1 to S3
MDAR Reproducibility Checklist
Movies S1 to S4

View/request a protocol for this paper from *Bio-protocol*.

Submitted 22 August 2022; resubmitted 30 March 2023
Accepted 17 May 2023
10.1126/science.ade5308

WORKING LIFE

*By* **Brita Kilburg-Basnyat**

# It's not just about the money

"I don't think industry would be a good fit for me," the student said, "I'm not money-motivated." As an industry scientist myself, I struggled to hide my surprise. I was participating in a conference event to help graduate students and postdocs explore career options by connecting with professionals across industry, academia, government, and nonprofits. I love my work, and I was excited to share it with the attendees. Overall, they were enthusiastic to learn about their options, and some were seriously considering industry. I was glad to help them figure out whether it was right for them and, if so, how they could get there. So that student's dismissive comment caught me off-guard—though I suspect others felt the same way. After pausing to regain my composure, I shared my own journey.

When I was in graduate school, a relative I considered a surrogate grandmother and confidant, whom I spent holidays with and who even did a reading at my wedding, was diagnosed with late-stage ovarian cancer. She was hospitalized about an hour and a half away from her home and family—but near where I was living for graduate school. I visited her regularly and attended doctor's appointments with her. I put together a comfort purse for her chemotherapy treatment, complete with ginger chews, a blanket I had crocheted, and other items to help with side effects and pass the time. I was there to see the bruises on her arms from constant needles, to helplessly watch her become less and less herself as the chemotherapy ravaged her body and eventually stopped working. When she was transferred to an assisted living facility, I stopped by on my drive home to see her and read to her.

**"The majority of industry scientists ... value the contributions they make."**

During one visit, she said it was too bad my Ph.D. research wasn't about cancer. Instead, my doctorate was in toxicology: analyzing dust for allergens, bacteria, and fungi. Her comment struck me, but changing course didn't seem a viable option. My expertise felt too far from cancer. Plus, I had chosen to study toxicology because it is broadly important, marketable, and intellectually stimulating. Those reasons still seemed compelling, and the thought of abandoning my carefully laid plans because of an emotional personal experience felt unprofessional.

Eventually, she was transferred to hospice closer to her hometown, where she passed peacefully. And I went on to do my postdoc in a pharmacology and toxicology laboratory, assessing the impact of ozone exposure and dietary omega-3 fatty acids on pulmonary inflammation in mice.

It was during my postdoc that I first began to explore nonacademic career options, thanks in large part to my amazing, supportive adviser, who encouraged me to attend career development symposiums and helped me make connections and network. That's when pursuing drug development came on my radar. But I still hesitated to make the transition—until a graduate student asked me to proofread her award application. The essay explained that she wanted to get into drug development because of a loved one's illness. Her essay moved me and showed me a new way my career could have meaning.

In the 3.5 years I have been working in industry, I have found that meaning. The transition from academia has been an adjustment and a learning journey, but industry is a great fit for me. I now work for a company developing cancer treatments, and I am proud of the contributions I am making to help patients. I think of the dear relative I lost often and for me, that is motivating.

Contrary to what the student at the conference assumed, working in industry isn't all about the money for me. Of course making a living is important, and there's nothing wrong with financial considerations playing a role in career decisions. But the majority of industry scientists I know aren't doing it only for the money; they enjoy their jobs and value the contributions they make. And many of them, like me, have come to this work with a deeply personal motivation. And there's nothing wrong with that either. ■

Brita Kilburg-Basnyat works in oncology drug development in Wisconsin. Send your career story to SciCareerEditor@aaas.org.